

TESI DI LAUREA MAGISTRALE  
UNIVERSITA' DEGLI STUDI DI PADOVA  
CORSO DI LAUREA IN BIOINGEGNERIA  
ANNO ACCADEMICO 2013/2014

**Confronto di metodi per l'apprendimento  
strutturale delle reti bayesiane nel caso di dati  
incompleti**



**Laureando:** Alessandro Neodo, 1056443

**Relatrice:** Dott.ssa Barbara Di Camillo

**Correlatore:** Dott. Francesco Sambo



*A mio padre, grande esempio di onestà,  
ed a mia madre, grande esempio di pazienza.*



# Indice

<b>Indice</b>	<b>ii</b>
<b>Lista dei simboli</b>	<b>iii</b>
<b>Elenco delle tabelle</b>	<b>v</b>
<b>Elenco delle figure</b>	<b>vii</b>
<b>Elenco degli algoritmi</b>	<b>ix</b>
<b>Introduzione</b>	<b>1</b>
<b>1 Teoria</b>	<b>3</b>
1.1 Nozioni basilari . . . . .	3
1.1.1 Teoria delle probabilità . . . . .	3
1.1.2 Grafi . . . . .	13
1.2 Reti bayesiane . . . . .	16
<b>2 Apprendimento con dati incompleti</b>	<b>25</b>
2.1 Apprendimento di reti bayesiane . . . . .	25
2.1.1 Il problema della missingness . . . . .	26
2.1.2 Apprendimento di parametri . . . . .	27
2.1.3 Apprendimento di strutture . . . . .	33
2.2 Stato dell'arte . . . . .	36
2.2.1 Metodologie generiche . . . . .	37
2.2.2 MMHC (Max Min Hill Climbing) . . . . .	39
2.2.3 SEM (Structural Expectation Maximization) . . . . .	42
2.2.4 MWST-EM (Maximum Weight Spanning Tree - Expectation Maximization) . . . . .	43
2.2.5 GES-EM (Greedy Equivalence Search - Expectation Maximization) . . . . .	44
2.2.6 Ulteriori algoritmi . . . . .	46
<b>3 Comparazione</b>	<b>49</b>
3.1 Setup sperimentale . . . . .	49
3.1.1 Algoritmi . . . . .	50
3.1.2 Reti . . . . .	51
3.1.3 Dataset . . . . .	55
3.1.4 Misure di performance . . . . .	57

3.1.5	Prove . . . . .	59
3.2	Risultati . . . . .	60
3.2.1	Tabelle . . . . .	63
3.2.2	Grafici . . . . .	154
	<b>Conclusione</b>	<b>237</b>
	<b>Ringraziamenti</b>	<b>241</b>
	<b>Bibliografia</b>	<b>246</b>
	<b>Indice analitico</b>	<b>247</b>

# Lista dei simboli e delle abbreviazioni

Tabella 1: Lista dei simboli e delle abbreviazioni

<i>Simbolo</i>	<i>Spiegazione</i>
$\Omega$	Spazio campionario
$A$	Evento
$\mathbb{A}$	Spazio degli eventi
v.a.	Abbr. per <i>Variabile Aleatoria</i>
$X$	Variabile aleatoria
$x$	Valore associato alla v.a.
$P(X = x)$	Probabilità che $X$ sia uguale ad $x$
$P_X$	Funzione di distribuzione di $X$
$p_X$	Densità di probabilità di $X$
$E[X]$	Valore atteso della v.a. $X$
$var(X)$	Varianza della v.a. $X$
$\sigma_X$	Deviazione standard della v.a. $X$
$I_P(A, B   C)$	Gli insiemi $A$ e $B$ sono condizionatamente indipendenti dato $C$
$\mathcal{G}$	Grafo
$\mathcal{V}$	Insieme dei vertici di un grafo
$\mathcal{E}$	Insieme degli archi di un grafo
$u, v$	Vertici di un grafo
$(u, v)$	Arco (orientato o no) tra $u$ e $v$
$\mathcal{T}$	Albero
DAG	Sigla per <i>grafo diretto aciclico</i>
$PA_X$	Insieme dei padri del nodo $X$
$ND_X$	Insieme dei non-discendenti del nodo $X$
$P$	Distribuzione di probabilità di una rete bayesiana
$(\mathcal{G}, P)$	Rete bayesiana
$pa_i$	$i$ -esimo padre di un nodo
$nd_i$	$i$ -esimo non-discendente di un nodo

Tabella 2: Lista dei simboli e delle abbreviazioni (*continua*)

<i>Simbolo</i>	<i>Spiegazione</i>
$I_{\mathcal{G}}(X, Y   Z)$	$X$ e $Y$ d-separati da $Z$ in $\mathcal{G}$
BN	Abbr. per <i>rete bayesiana</i> (Bayesian Network)
$\mathcal{D}$	Dataset $\mathcal{D}$
$\mathcal{M}$	Matrice indicatrice dei dati mancanti
MAR	Abbr. per <i>Missing At Random</i>
MCAR	Abbr. per <i>Missing Completely At Random</i>
NMAR	Abbr. per <i>Not Missing At Random</i>
CCA	Abbr. per <i>Complete Cases Analysis</i>
ACA	Abbr. per <i>Available Cases Analysis</i>

# Elenco delle tabelle

1	Lista dei simboli e delle abbreviazioni . . . . .	iii
2	Lista dei simboli e delle abbreviazioni ( <i>continua</i> ) . . . . .	iv
3.1	Caratteristiche delle reti utilizzate nel nostro studio. . . . .	55



# Elenco delle figure

1.1	Esempio di rete bayesiana (da (Russell e Norvig 2003)). . . . .	18
1.2	Esempio di Markov Blanket. Da (Riggelsen 2006). . . . .	22



# List of Algorithms

1	Algoritmo MMPC0 . . . . .	40
2	Algoritmo MMPC . . . . .	41
3	Algoritmo MMHC . . . . .	41
4	Algoritmo AMS-EM . . . . .	42
5	Algoritmo SEM-Bayesiano . . . . .	43
6	Algoritmo MWST-EM . . . . .	44
7	Algoritmo GES-EM - Inserzione . . . . .	46



# Introduzione

Le reti bayesiane (Bayesian Networks, BN) costituiscono un potente mezzo per lo studio dei legami che intercorrono tra le variabili di un insieme di dati. Tale tecnica è spesso utilizzata per modellare domini di interesse, fornire supporto alle decisioni e trovare strategie di risoluzione dei problemi nell'ambito di domini con incertezza. Gli ambiti di applicazione sono i più disparati, spaziando dalla meteorologia (Jensen e Jensen 1996) all'economia (Binder et al. 1997), dall'agricoltura (Kristensen e Rasmussen 2002) fino alla medicina (Andreassen et al. 1991) (Onisko 2003).

Una rete bayesiana è una coppia  $(\mathcal{G}, P)$ , con  $\mathcal{G}$  grafo diretto aciclico (Directed Acyclic Graph, DAG) e  $P$  distribuzione di probabilità, per cui valga una specifica proprietà, detta *condizione di Markov* (per la spiegazione si rimanda al cap. 1). Sotto un'ulteriore ipotesi (*faithfulness*), nella pratica spesso verificata o quantomeno assunta, gli archi del grafo rappresentano legami di causalità diretta tra le variabili in gioco. Il duplice vantaggio delle reti bayesiane sta proprio nella rappresentazione grafica dei legami di causalità, immediatamente comprensibile anche ai non addetti ai lavori, e nella possibilità di eseguire inferenze statistiche sul verificarsi di un evento non osservabile in modo efficiente, utilizzando metodi che si ispirano alla legge di Bayes.

La grande sfida aperta di tale approccio risiede nella necessità di imparare la distribuzione  $P$  (i cui elementi sono definiti *parametri*) e/o il grafo (definito *struttura*) a partire dai dati. Purtroppo, mentre l'apprendimento parametrico (una volta nota la struttura, ad esempio perché decretata da un esperto) è relativamente semplice ed efficace, imparare il DAG è un'operazione notevolmente più complessa. In particolare, molti metodi richiedono che i dataset da cui partire per apprendere la struttura siano completi, ovvero che non abbiano alcun dato mancante. Evidentemente, questa è una richiesta che in certi contesti, come può essere quello biomedico, appare alquanto irrealistica e compromette notevolmente la possibilità di applicare le reti bayesiane a casi reali.

Allo stato attuale, esistono pochi algoritmi che affrontano il problema dell'apprendimento di strutture di rete a partire da dati incompleti. Inoltre, al meglio delle nostre conoscenze, non esiste una comparazione tra i metodi che sia vasta e terza. L'unica cosa che sembra possibile affermare è che il gold standard *de facto* sia dato dall'algoritmo Structural Expectation Maximization proposto da Nir Friedman in (Friedman 1997) (Friedman 1998), che però ha quasi vent'anni. Non ci risulta che nel frattempo siano usciti molti metodi nuovi con performance significativamente migliori e pare che nessuno di questi abbia mai raggiunto grande popolarità.

L'obiettivo di questa tesi è pertanto di eseguire una comparazione estesa tra alcuni algoritmi e strategie per l'apprendimento strutturale di reti bayesiane a

partire da dati incompleti. L'approccio seguito consiste nel generare dataset partendo da reti note presenti in letteratura (che costituiranno il nostro gold standard), simulare la presenza di dati mancanti secondo diversi paradigmi e percentuale di incompletezza, ricostruirle con i metodi in esame e utilizzare alcune metriche per valutare la qualità della ricostruzione del grafo. Poiché, una volta ricostruita correttamente la struttura di una rete, l'apprendimento parametrico è facile e, in un certo senso, "standard", ci è sembrato più corretto ed interessante valutare soltanto l'aspetto strutturale. Una certa attenzione è stata infine posta alla valutazione dell'uso combinato di più metodi, onde cercare di comprendere se esistano algoritmi che danno risultati complementari o che comunque, usati assieme, migliorino le prestazioni rispetto a ciascuno dei singoli.

I risultati sperimentali indicano che non sembra esistere un metodo spiccatamente migliore degli altri. In particolare, gli approcci che si sono rivelati più performanti secondo i criteri adottati (Structural Hamming Distance, definita in (Tsamardinos, Brown e Aliferis 2006), precisione, recall, tempo), sono stati SEM+T (SEM inizializzato con una rete ottenuta da MWST-EM) (Leray e Francois 2005) e k-nn seguito da MMHC (Tsamardinos, Brown e Aliferis 2006), anche se buoni risultati sono stati ottenuti anche da altri metodi (ad esempio, da SEM inizializzato con rete vuota). L'ensemble learning ci ha invece permesso di concludere che i due approcci già citati (SEM+T e k-nn con MMHC) sembrano essere preferibili agli altri, nonostante anche qui la variabilità di risultati sia molto alta. In particolare, l'uso combinato dei due algoritmi migliora sempre in affidabilità rispetto al metodo peggiore mantenendosi però (anche se non necessariamente) inferiore alle performance del migliore. Tenuto conto che in un contesto reale non si sa quale dei due garantisca migliore qualità, l'uso congiunto può essere più che ragionevole, se non consigliato.

Questa tesi è suddivisa in tre parti:

**Teoria** In questo capitolo vengono presentate le principali nozioni teoriche riguardanti le reti bayesiane. Si parte con una presentazione delle basi di Teoria delle Probabilità e Teoria dei Grafi, per poi proseguire con definizioni e teoremi necessari per presentare le reti bayesiane.

**Apprendimento con dati incompleti** In questo capitolo sono illustrati gli algoritmi utilizzati per la comparazione. Si presenta il problema dell'apprendimento di BN sia per quanto riguarda i parametri che per ciò che concerne l'apprendimento di struttura. Infine, ci si concentra sulla descrizione degli algoritmi utilizzati nella tesi.

**Comparazione** Qui si descrive l'impianto sperimentale ed i risultati ottenuti. Verranno illustrate le varianti degli algoritmi utilizzate, le reti ed i dataset, i tipi di prove utilizzate e le misure di performance a cui si è fatto ricorso (Structural Hamming Distance, precisione, recall, tempo, area sotto la curva delle curve Precision vs. Recall). Infine, verranno riportati tutte le tabelle ed i grafici ottenuti.

# Capitolo 1

## Teoria

In questo capitolo verranno introdotte le principali nozioni teoriche necessarie per la comprensione dei concetti utilizzati nel proseguio di questa tesi (nozioni basilari su teoria delle probabilità e grafi, reti bayesiane e apprendimento di reti). Data la vastità degli argomenti, non c'è, ovviamente, nessuna pretesa di essere esaustivi e per una trattazione più approfondita si rimanda il lettore alla bibliografia. Per quanto riguarda la simbologia adottata, i simboli e le abbreviazioni saranno spiegati la prima volta che compariranno, ma è comunque presente una tabella riassuntiva all'inizio dell'opera.

### 1.1 Nozioni basilari

Una rete bayesiana è composta da due parti: un grafo (diretto aciclico) e una distribuzione di probabilità (dotata di alcune particolari caratteristiche che verranno descritte in seguito). Per prima cosa, pertanto, è necessario introdurre alcune nozioni basilari di teoria delle probabilità e grafi.

#### 1.1.1 Teoria delle probabilità

La Teoria delle Probabilità è quella branca della Matematica che si prefigge di studiare e descrivere l'incertezza con cui è possibile fare affermazioni su determinati eventi. Esistono, per ragioni storiche e filosofiche, almeno tre diversi approcci con cui la nozione di probabilità viene concepita ed usata: *ratio*, frequentista e soggettiva. Per quanto riguarda la formalizzazione matematica di questo concetto, invece, gli assiomi alla base della teoria sono dovuti ad A. N. Kolmogorov.

Se non diversamente indicato, le definizioni ed i teoremi riportati in questa sezione sono adattati da (Bertsekas e Tsitsikilis 2008), (Baldi 1998) e (Neapolitan 2003).

#### Modello probabilistico

La descrizione matematica di una situazione di incertezza (altresì chiamata *esperienza*) richiede un modello probabilistico. Iniziamo dando alcune definizioni elementari:

**Definizione 1.1** (Spazio campionario). Si definisce *spazio campionario* l'insieme  $\Omega$  di tutti i possibili esiti di un esperimento.

**Definizione 1.2** (Evento). Un *evento* è un qualsiasi sottoinsieme  $A \subseteq \Omega$ . Lo *spazio degli eventi* è l'insieme  $\mathbb{A}$  di tutti gli eventi di  $\Omega$ .

Un evento  $A$  si dice *elementare* se contiene un solo elemento.

Due eventi  $A$  e  $B$  si diranno *disgiunti* se  $A \cap B = \emptyset$ .

Si noti come la definizione includa tra gli eventi anche l'insieme vuoto  $\emptyset$  e l'intero spazio campionario  $\Omega$ .

Possiamo quindi dare la definizione di modello probabilistico:

**Definizione 1.3** (Modello probabilistico). Diremo *modello probabilistico* l'insieme di tre elementi  $(\Omega, \mathbb{A}, P)$ :

1. Uno *spazio campionario*  $\Omega$ .
2. Uno *spazio degli eventi*  $\mathbb{A}$ .
3. Una *legge di probabilità*  $P$ , cioè una funzione che, per ogni evento  $A$ , restituisca un valore non-negativo  $P(A)$ , detto *probabilità di  $A$* , in maniera coerente con gli assiomi di Kolmogorov.

Un modello probabilistico viene anche detto *spazio di probabilità*.

Gli assiomi a cui una legge di probabilità deve sottostare sono:

**Definizione 1.4** (Assiomi di Kolmogorov). Siano dati uno spazio campionario  $\Omega$ , il relativo spazio degli eventi  $\mathbb{A}$  e due eventi  $A$  e  $B \in \mathbb{A}$ . La funzione  $P : \mathbb{A} \rightarrow \mathbb{R}$  è una legge di probabilità se valgono le seguenti condizioni:

1. (Non-negatività)  $P(A) \geq 0$ , per ogni  $A \in \mathbb{A}$
2. (Additività) Siano  $A$  e  $B$  due eventi disgiunti, allora  $P(A \cup B) = P(A) + P(B)$ . Più in generale, se lo spazio campionario ha infiniti elementi, allora una sequenza di eventi disgiunti  $A_1, A_2, \dots$  deve soddisfare

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

3. (Normalizzazione) La probabilità dell'intero spazio campionario  $\Omega$  è pari ad 1, cioè:  $P(\Omega) = 1$ .

Alcune semplici proprietà che derivano immediatamente dagli assiomi di Kolmogorov sono:

**Proposizione 1.1.** Siano dati una legge di probabilità  $P$  e due eventi  $A$  e  $B \subseteq \Omega$ .

1. L'insieme vuoto ha probabilità zero,  $P(\emptyset) = 0$ .
2. Se  $A \subset B$ , allora  $P(A) \leq P(B)$ .
3.  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ .
4.  $P(A \cap B) \leq P(A) + P(B)$ .

Dal punto di vista pratico, però, che cosa rappresenta la nozione di probabilità? Sostanzialmente, la si può ritenere una forma di misura della verosimiglianza con cui un evento può accadere. A tal proposito, esistono tre approcci con cui questa misura viene definita.

La prima si basa sul principio di indifferenza:

**Teorema 1.1** (Principio di indifferenza). *Sia dato uno spazio campionario  $\Omega$ . In assenza di motivi per cui preferire un evento all'altro, i suoi eventi elementari sono considerati equiprobabili.*

*Qualora lo spazio sia discreto con  $n$  elementi, ciò equivale a definire la probabilità di ogni evento elementare  $E$  come:*

$$P(E) = \frac{1}{n}$$

e, nel caso generale di un evento  $A$ ,

$$P(A) = \frac{|A|}{n}$$

Una probabilità definita in questo modo prende il nome di *ratio*. È questo il caso in cui si abbiano tre palline di tre colori diversi in un'urna e si voglia valutare la probabilità di estrarne una di un determinato colore.

Il secondo approccio è quello *frequentista*, in cui la probabilità viene assegnata sulla base della frequenza in cui si verifica un evento in esperimenti simili a quello in esame. Un classico esempio può essere quello del lancio di una moneta. In questo caso, ripetendo il lancio  $n$  volte, potremmo stimare la probabilità di avere testa come

$$P(\text{testa}) = \frac{\text{numero di volte in cui si ottiene testa}}{n}$$

In via teorica, andremo a definire la probabilità dell'evento  $A$  come

$$P(A) = \lim_{n \rightarrow \infty} \frac{\text{numero di volte in cui si verifica } A}{n}$$

dove  $n$  indica il numero di volte in cui l'esperimento è stato condotto (che, idealmente, tende all'infinito).

Infine, il terzo approccio è quello *soggettivo*. Le probabilità che una squadra vinca una partita o che l'Italia esca dalla recessione entro un anno, ad esempio, non possono essere calcolate nè come *ratio* nè con un approccio frequentista, in quanto, anche conoscendo il comportamento passato degli eventi, le condizioni sperimentali variano di volta in volta. Pertanto, dovendo esprimere una valutazione sul grado di incertezza su un evento di questo tipo, si è costretti ad attenersi a valutazioni soggettive.

### Probabilità condizionata

In certe situazioni si dispone di informazione parziale riguardo l'esito di un esperimento e questa conoscenza può alterare la probabilità che associamo ad un evento. Ad esempio, potremmo conoscere la somma del lancio di due dadi e

voler sapere con quale probabilità il primo dado vale 6. Se la somma valesse 2,  $P(\text{Dado}_1 = 6) = 0$ . D'altra parte, se la somma valesse 7,  $P(\text{Dado}_1 = 6) = 1/6$ . Si dà questa definizione formale:

**Definizione 1.5** (Probabilità condizionata). Siano dati due eventi  $A$  e  $B \subseteq \Omega$ . La *probabilità condizionata di  $A$  dato  $B$* , con  $P(B) > 0$  è data da:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (1.1)$$

e definisce una nuova legge di probabilità (condizionata) sullo stesso spazio campionario  $\Omega$ . In particolare, tutte le proprietà delle leggi di probabilità valgono anche per le leggi di probabilità condizionata.

Due importanti teoremi relativi alle probabilità condizionate sono la regola del prodotto ed il teorema della probabilità totale.

**Teorema 1.2** (Regola del prodotto). *Se tutti gli eventi condizionanti hanno probabilità positiva, allora*

$$P(\cap_{i=1}^n A_i) = P(A_1)P(A_2 | A_1)P(A_3 | A_2 \cap A_1) \dots P(A_n | \cap_{i=1}^{n-1} A_i) \quad (1.2)$$

**Teorema 1.3** (della probabilità totale). *Siano  $A_1 \dots A_n$   $n$  eventi disgiunti che costituiscono una partizione dello spazio campionario  $\Omega$ , cioè tali che ogni  $A \subseteq \Omega$  sia contenuto in uno ed un solo evento  $A_i$ . Supponendo che  $P(A_i) > 0 \quad \forall i$ , allora, dato un qualsiasi evento  $B$ , si ha:*

$$\begin{aligned} P(B) &= P(A_1 \cap B) + \dots + P(A_n \cap B) \\ &= P(A_1)P(B | A_1) + \dots + P(A_n)P(B | A_n) \end{aligned} \quad (1.3)$$

L'importanza di questi teoremi sta nel fatto che ci consentono di valutare la probabilità di un qualsiasi evento una volta che siano note le probabilità condizionate che lo legano ad un ulteriore insieme di eventi. Si tratta, in un certo senso, di un approccio *divide et impera* per il calcolo delle probabilità.

Tra i teoremi più importanti relativi alle probabilità condizionate (e certamente il teorema fondamentale per questa tesi) c'è la legge di Bayes, tipicamente usata in associazione al teorema della probabilità totale.

**Teorema 1.4** (Legge di Bayes). *Siano  $A_1 \dots A_n$   $n$  eventi disgiunti che costituiscono una partizione dello spazio campionario  $\Omega$  tali che  $P(A_i) > 0 \quad \forall i$ . Allora, dato un qualsiasi evento  $B$  tale che  $P(B) > 0$ , si ha:*

$$\begin{aligned} P(A_i | B) &= \frac{P(A_i)P(B | A_i)}{P(B)} \\ &= \frac{P(A_i)P(B | A_i)}{P(A_1)P(B | A_1) + \dots + P(A_n)P(B | A_n)} \end{aligned} \quad (1.4)$$

### Indipendenza

Un caso di particolare interesse nell'ambito delle probabilità condizionate si ha quando essere a conoscenza che si è verificato un certo evento  $B$  non altera in alcun modo la probabilità che assegnamo ad un altro evento  $A$ . In altre parole, la probabilità condizionata  $P(A | B)$  è uguale alla probabilità iniziale  $P(A)$ . Si dice in questo caso che i due eventi sono indipendenti.

**Definizione 1.6** (Eventi indipendenti). Due eventi  $A$  e  $B$  sono *indipendenti* se

$$P(A \cap B) = P(A)P(B) \quad (1.5)$$

Inoltre, se  $P(B) > 0$ , dalla definizione di probabilità condizionata si ha

$$P(A | B) = P(A) \quad (1.6)$$

Una definizione analoga si può dare nel caso di probabilità condizionate:

**Definizione 1.7** (Indipendenza condizionata). Siano dati tre eventi  $A$ ,  $B$  e  $C$ , con  $P(C) > 0$ .  $A$  e  $B$  sono legati da *indipendenza condizionata rispetto a  $C$*  se

$$P(A \cap B | C) = P(A | C)P(B | C) \quad (1.7)$$

Inoltre, se  $P(B \cap C) > 0$ , dalla definizione di probabilità condizionata si ha

$$P(A | B \cap C) = P(A | C) \quad (1.8)$$

In maniera naturale si può estendere la definizione di indipendenza al caso di eventi multipli:

**Definizione 1.8** (Indipendenza di un insieme di eventi). Gli eventi  $A_1, \dots, A_n$  sono *indipendenti* se

$$P(\cap_{i \in S} A_i) = \prod_{i \in S} P(A_i), \quad \forall S \subseteq \{1, 2, \dots, n\} \quad (1.9)$$

Ad esempio, nel caso di tre eventi  $A_1, A_2$  e  $A_3$ , l'indipendenza implica:

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1)P(A_2) \\ P(A_1 \cap A_3) &= P(A_1)P(A_3) \\ P(A_2 \cap A_3) &= P(A_2)P(A_3) \\ P(A_1 \cap A_2 \cap A_3) &= P(A_1)P(A_2)P(A_3) \end{aligned}$$

È importante evidenziare alcune osservazioni a proposito dell'indipendenza.

**Proposizione 1.2.** 1. Se due eventi  $A$  e  $B$  sono indipendenti, allora anche  $A$  e  $B^C$  lo sono.

2. L'indipendenza condizionata non implica l'indipendenza, nè vale il viceversa.

3. Il fatto che un insieme di eventi siano a due a due indipendenti non implica che valga l'indipendenza multipla. Vale invece il viceversa, com'è ovvio dalla definizione.

## Variabili aleatorie

Spesso si è interessati a calcolare probabilità per quantità che sono funzioni del risultato di un fenomeno casuale. Supponiamo di lanciare una moneta dieci volte, puntando 1 euro ad ogni lancio e vincendone 1.5 se esce testa, 0 altrimenti. In questo caso, più che essere interessati alla probabilità che, per esempio, escano 6 teste e 4 croci, ci chiederemo con quale probabilità alla fine dei lanci il nostro capitale sia cresciuto, oppure se, mediamente, ci aspettiamo che il capitale cresca o cali. Per rispondere a queste domande, si introduce il concetto di variabile aleatoria.

**Definizione 1.9** (Variabile aleatoria). Dato uno spazio di probabilità  $(\Omega, \mathbb{A}, P)$ , una *variabile aleatoria* (v.a.)  $X$  è una funzione  $X : \Omega \rightarrow \mathbb{R}$  tale che, per ogni  $t \in \mathbb{R}$ , l'insieme  $\{\omega \in \Omega \mid t.c. \ X(\omega) \leq t\} \in \mathbb{A}$ . L'insieme dei valori che  $X$  può assumere è chiamato *spazio della variabile aleatoria*  $X$ .

In altre parole, si chiede alla v.a. di essere una funzione per cui si possa calcolare  $P(\{\omega \in \Omega \mid t.c. \ X(\omega) \leq t\})$ , che indica la probabilità che  $X$  assuma valori inferiori a  $t$ . Nel seguito, per semplicità di notazione, ometteremo la dipendenza di  $X$  dagli eventi  $\omega$  e scriveremo quindi, ad esempio,  $P\{X \leq t\}$ . E' importante far presente che una funzione di v.a. è anch'essa una variabile aleatoria.

Esistono due categorie di variabili aleatorie, in base ai valori che la variabile può assumere: se il suo spazio è finito o al più un infinito numerabile, la v.a. viene definita *discreta*, altrimenti si dice *continua*. In questa tesi si utilizzeranno soltanto v.a. discrete e quindi la trattazione si limiterà ad esse. Il caso continuo ne è una generalizzazione e le definizioni che seguono possono sostanzialmente essere ripetute sostituendo le sommatorie con degli opportuni integrali. In ogni caso, non ci soffermeremo oltre sull'argomento e si rimanda il lettore interessato a (Bertsekas e Tsitsikilis 2008).

Nel definire una variabile aleatoria, siamo interessati alle probabilità che i valori del suo spazio possono assumere. Questo concetto è espresso mediante la funzione (o legge) di distribuzione e la densità di probabilità, definite come segue.

**Definizione 1.10** (Densità di probabilità discreta). Data una v.a. discreta  $X$ , una funzione  $p_X : \mathbb{R} \rightarrow \mathbb{R}^+$  si dice *densità di probabilità discreta* se

1.  $p_X(x) = 0$  tranne al più per una infinità numerabile di valori (assunti da  $X$ );

2.

$$\sum_{x \in \mathbb{R}} p_X(x) = 1 \quad (1.10)$$

In particolare, essa sarà definita come  $p_X(x) = P(\{X = x\})$ .

**Definizione 1.11** (Funzione di distribuzione discreta). Siano dati uno spazio di probabilità  $(\Omega, \mathbb{A}, P)$  ed una v.a.  $X$  su di esso definita. La *funzione di distribuzione di  $X$*  è una funzione  $P_X : \mathbb{R} \rightarrow [0, 1]$  tale che, per ogni  $S \subseteq \mathbb{R}$  restituisce la probabilità che  $X$  assuma valori in  $S$ , cioè

$$P_X(\{S\}) = P(\{X \in S\}) \quad (1.11)$$

In pratica, per comodità, non ci riferiremo mai direttamente a  $P_X$ , ma continueremo ad usare la funzione di probabilità nelle forme semplificate  $P(X = x) = P(\{X = x\})$  e  $P(X \in S) = P(\{X \in S\})$ .

Nel caso discreto, dall'assioma di additività deriva un immediato legame tra funzione di distribuzione e densità di probabilità:

$$P(X \in S) = \sum_{x \in S} P(X = x) = \sum_{x \in S} p_X(x) \quad (1.12)$$

Un'altra importante funzione è la funzione di ripartizione.

**Definizione 1.12** (Funzione di ripartizione). Data una v.a.  $X$ , si chiama *funzione di ripartizione* la funzione  $F_X : \mathbb{R} \rightarrow [0, 1]$  definita da

$$F_X(t) = P(X \leq t)$$

Si tratta di una funzione non-negativa e non-decrescente, la cui conoscenza è equivalente a quella della distribuzione di  $X$ , infatti da 1.12

$$F_X(t) = \sum_{x \leq t} p_X(x)$$

Per riassumere le informazioni contenute nella distribuzione di probabilità, è spesso conveniente l'uso di due quantità: il valore atteso (detto anche media o speranza matematica), che rappresenta una media pesata secondo le probabilità della v.a., e la varianza, che quantifica la dispersione rispetto alla media. Se ne forniscono le definizioni.

**Definizione 1.13** (Valore atteso). Data una v.a.  $X$  con densità  $p_X$ , il suo *valore atteso* è dato da

$$E[X] = \sum_x x p_X(x) \quad (1.13)$$

Più in generale, si dà la definizione di momento di ordine  $n$ .

**Definizione 1.14** (Momento). Data una v.a.  $X$  con densità  $p_X$  ed un valore  $n \in \mathbb{N}$ , il suo *momento di ordine  $n$*  è dato da

$$E[X^n] = \sum_x x^n p_X(x) \quad (1.14)$$

**Definizione 1.15** (Varianza). Data una v.a.  $X$  con densità  $p_X$ , la sua *varianza* è data da

$$\begin{aligned} \text{var}(X) &= E[(X - E[X])^2] \\ &= \sum_x (x - E[X])^2 p_X(x) \\ &= E[X^2] - (E[X])^2 \end{aligned} \quad (1.15)$$

Si tratta di un valore sempre non-negativo. La sua radice quadrata è detta *deviazione standard* e si indica con  $\sigma_X$ .

Funzioni di variabili aleatorie sono anch'esse variabili aleatorie e se ne può calcolare densità e valore atteso con facilità.

**Proposizione 1.3** (Densità e valore atteso funzioni di v.a.). *Siano date due v.a.  $X$  e  $Y$  legate da una funzione  $g : X \rightarrow Y$ , ovvero valga*

$$Y = g(X)$$

*allora si ha che*

$$p_Y(y) = \sum_{\{x|g(x)=y\}} p_X(x)$$

*e*

$$E[g(X)] = \sum_x g(x) p_X(x)$$

Nel caso particolare di funzioni lineari di v.a. si ha il seguente risultato.

**Proposizione 1.4** (Media e varianza di funzioni lineari di v.a.). *Sia  $X$  una v.a. ed  $Y$  definita come*

$$Y = aX + b$$

Allora valgono le seguenti relazioni:

$$E[Y] = aE[X] + b \quad (1.16)$$

e

$$\text{var}(Y) = a^2 \text{var}(X) \quad (1.17)$$

### Probabilità congiunta

Nel caso di funzioni a valori in  $\mathbb{R}^n$  la nozione di v.a. viene estesa come segue.

**Definizione 1.16** (Variabile aleatoria n-dimensionale). Una v.a. n-dimensionale discreta è una funzione  $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$  tale che le funzioni  $X_1, \dots, X_n$  siano v.a. reali discrete.

Si estende anche la nozione di densità di probabilità.

**Definizione 1.17** (Densità congiunta). Data una v.a. discreta n-dimensionale  $X = (X_1, \dots, X_n)$ , la funzione  $p_X : \mathbb{R}^n \rightarrow \mathbb{R}^+$ , definita come  $p_X(x) = P(\{X = x\}) = P(X_1 = x_1, \dots, X_n = x_n)$ , si dice *densità di probabilità congiunta di  $X$*  se

1.  $p_X(x) \geq 0$  per ogni  $x \in \mathbb{R}^n$ ;
2.  $p_X(x) = 0$  tranne al più per una infinità numerabile di valori (assunti da  $X$ );
- 3.

$$\sum_{x \in \mathbb{R}^n} p_X(x) = 1 \quad (1.18)$$

Le densità  $p_1, \dots, p_n$  delle singole v.a. reali  $X_1, \dots, X_n$  si chiamano invece *densità marginali di  $X$*  e si calcolano a partire dalla densità congiunta come

$$p_{X_i}(x) = \sum_{X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n} p_X(x) \quad (1.19)$$

Non è invece in generale possibile risalire alla densità congiunta partendo dalle marginali, perchè possono essere comuni a diverse densità congiunte. Anche in questo caso funzioni di v.a. sono v.a. e la proposizione 1.3 si estende in modo naturale.

### Densità condizionate

Anche il concetto di probabilità condizionata è estendibile al caso di variabili aleatorie, solo che in questo caso il condizionamento può avvenire sia rispetto ad un evento, sia rispetto ad un'altra v.a.

**Definizione 1.18** (Densità condizionata). Siano date due variabili aleatorie  $X$  e  $Y$  ed un evento  $A$ , tutti associati allo stesso esperimento.

La densità condizionata di  $X$  rispetto all'evento  $A$ , con  $P(A) > 0$  è data da

$$p_{X|A}(x) = P(X = x | A) = \frac{P(\{X = x\} \cap A)}{P(A)}$$

ed è tale che

$$\sum_x p_{X|A}(x) = 1$$

La densità condizionata di  $X$  rispetto alla v.a.  $Y$ , con  $P(Y = y) > 0$  per un  $y$  fissato, è data da

$$\begin{aligned} p_{X|Y}(x | y) &= P(X = x | Y = y) \\ &= \frac{P(\{X = x, Y = y\})}{P(Y = y)} \\ &= \frac{p_{X,Y}(\{x, y\})}{p_Y(y)} \end{aligned}$$

ed è tale che

$$\sum_x p_{X|Y}(x | y) = 1$$

L'estensione al caso di più di due variabili aleatorie è immediata.

In entrambi i casi è possibile ricavare la densità di  $X$  note le probabilità condizionate, secondo quanto indicato nella seguente proposizione.

**Proposizione 1.5.** Sia  $A_1, \dots, A_n$  una partizione dello spazio campionario, con  $P(A_i) > 0$  per ogni  $i$ . Allora vale

$$p_X(x) = \sum_{i=1}^n P(A_i) p_{X|A_i}(x)$$

Inoltre, per ogni evento  $B$ , con  $P(A_i \cap B) \geq 0$ , si ha

$$p_{X|B}(x) = \sum_{i=1}^n P(A_i | B) p_{X|A_i \cap B}(x)$$

Nel caso di due v.a.  $X$  e  $Y$  indipendenti, la densità marginale può essere calcolata con la formula

$$p_X(x) = \sum_y p_Y(y) p_{X|Y}(x | y)$$

**Teorema 1.5** (Regola della catena per v.a.). Siano date  $n$  variabili aleatorie  $X_1, \dots, X_n$  definite sullo stesso spazio campionario  $\Omega$ . Se  $P(x_1, x_2, \dots, x_n) \neq 0$ , allora

$$P(x_1, x_2, \dots, x_n) = P(x_n | x_{n-1}, x_{n-2}, \dots, x_1) \dots P(x_2 | x_1) P(x_1)$$

La legge di Bayes vale anche nel caso delle v.a. ed esiste in molteplici varianti che includono anche i casi di v.a. continue. Qui si riporta soltanto il caso di due variabili aleatorie discrete.

**Teorema 1.6** (Legge di Bayes per v.a. discrete). *Siano date due v.a. discrete  $X$  e  $Y$ . Allora vale*

$$p_{Y|X}(y|x) = \frac{p_Y(y)p_{X|Y}(x|y)}{\sum_{y'} p_Y(y')(x, y')} \quad (1.20)$$

### Indipendenza di variabili aleatorie

È possibile dare una definizione di indipendenza anche per le variabili aleatorie.

**Definizione 1.19** (Indipendenza di una v.a. da un evento). La v.a. (discreta)  $X$  e l'evento  $A$  definiti su uno stesso spazio di probabilità si dicono indipendenti se e solo se per ogni  $x \in X$  si ha

$$P(X = x \text{ e } A) = P(X = x)P(A) \quad (1.21)$$

**Definizione 1.20** (Indipendenza di variabili aleatorie). Le v.a. (discrete)  $X_1, \dots, X_n$  definite su uno stesso spazio di probabilità si dicono indipendenti se e solo se per ogni scelta di  $S_1, \dots, S_n \subset \mathbb{R}$  si ha

$$P(X_1 \in S_1, \dots, X_m \in S_m) = P(X_1 \in S_1)P(X_2 \in S_2) \dots P(X_n \in S_n) \quad (1.22)$$

Diremo che le infinite v.a.  $X_1, \dots, X_n, \dots$  sono indipendenti se e solo se per ogni  $m \geq 0$  risultano tra loro indipendenti le v.a.  $X_1, \dots, X_m$

A questo punto, una definizione importante per le reti bayesiane è quella di indipendenza condizionata.

**Definizione 1.21** (Indipendenza condizionata). Sia dato uno spazio di probabilità  $(\Omega, \mathbb{A}, P)$  e tre insiemi  $A$ ,  $B$  e  $C$  contenenti v.a. definite in  $\Omega$ . Diremo che gli insiemi  $A$  e  $B$  sono legati da *indipendenza condizionata rispetto a  $C$*  se, per tutti i valori  $a$ ,  $b$  e  $c$  contenuti negli insiemi, ogni volta che  $P(c) \neq 0$ , gli eventi  $A = a$  e  $B = b$  sono legati da indipendenza condizionata rispetto all'evento  $C = c$ , ovvero vale una delle seguenti condizioni:

- $P(a|c) = 0$
- $P(b|c) = 0$
- $P(a|b, c) = P(a|c)$

In questo caso, scriveremo

$$I_p(A, B | C)$$

### Inferenza bayesiana

L'Inferenza Statistica differisce dalla Teoria delle Probabilità in quanto in quest'ultima si assume che un determinato modello ubbidisca a degli assiomi e questo porta a poter ricavare un'unica risposta corretta. In Statistica, d'altro canto, per ogni problema possono esserci più metodi sensati e le risposte non essere univoche. In generale non c'è il miglior metodo a priori, salvo imporre vincoli addizionali o fare specifiche assunzioni. Anzi, la scelta di un metodo piuttosto che un altro è tipicamente legata a fattori diversi, quali migliori risultati e consenso

della comunità statistica sull'applicabilità di un metodo ad una determinata classe di problemi, ma anche esperienza e buon senso.

Le due principali scuole di pensiero sono quella classica -secondo cui i modelli incogniti sono quantità deterministiche sconosciute- e quella bayesiana -in cui i modelli incogniti sono trattati come variabili aleatorie con distribuzione nota-. In questa tesi siamo interessati all'approccio bayesiano.

Nella scuola bayesiana, si considera che il modello incognito sia scelto casualmente da una classe di modelli. Per far ciò, esso viene caratterizzato con una v.a.  $\Theta$  di cui viene assunta una densità di probabilità (nota)  $p_{\Theta}(\theta)$  definita *prior*. A questo punto, a partire dall'osservazione di variabili aleatorie correlate  $X = (X_1, \dots, X_n)$  chiamate *osservazioni* o *misure*, è possibile costruire la distribuzione condizionata  $p_{X|\Theta}$  (supponendo nota la distribuzione congiunta  $p_{X,\Theta}$ ), la quale costituisce il nostro modello del vettore delle osservazioni  $X$ . Infine, una volta che si sia eseguita una specifica osservazione  $x$ , mediante la legge di Bayes 1.6, è possibile ottenere la risposta desiderata, cioè la *distribuzione a posteriori* o *posterior*  $p_{\Theta|x}$  di  $\Theta$ .

Il motivo per cui si necessita dell'utilizzo della legge di Bayes e non si calcola direttamente la posterior a partire dallo spazio campionario è che nelle applicazioni non procediamo nell'ordine descritto in precedenza (definizione dello spazio campionario, determinazione delle probabilità degli eventi elementari e successivamente di variabili aleatorie e distribuzioni congiunte), ma partiamo direttamente dalle v.a. e cerchiamo relazioni di probabilità tra di esse, tenendo conto che spesso le quantità di interesse non sono quelle direttamente accessibili.

### 1.1.2 Grafi

Se non diversamente indicato, il contenuto di questa sezione è adattato da (Goodrich e Tamassia 2006) e (Cormen et al. 2005).

**Definizione 1.22** (Grafo). Un *grafo*  $\mathcal{G}$  è una coppia  $(\mathcal{V}, \mathcal{E})$ , in cui  $\mathcal{V}$  è un insieme finito e non vuoto detto *insieme dei vertici di  $G$*  (ed i suoi elementi sono detti *vertici* o *nodi*) e  $\mathcal{E}$ , chiamato *insieme degli archi*, è un insieme di coppie di vertici di  $\mathcal{V}$ . Se una coppia  $(u, v)$  è ordinata con  $u$  che precede  $v$  l'arco si dice *orientato*, altrimenti  $(u, v)$  indicherà lo stesso arco di  $(v, u)$  e sarà definito *non orientato*. Un grafo in cui tutti gli archi siano orientati sarà *diretto* (o *orientato*), uno in cui tutti gli archi siano non orientati sarà detto *non diretto* (o *non orientato*). Qualora  $\mathcal{G}$  abbia entrambi i tipi di arco, lo si definirà *grafo misto*. Si noti che ogni arco non orientato può essere convertito in una coppia di archi orientati. Diremo *sottografo* di  $\mathcal{G}$  un grafo  $\mathcal{H}$  i cui vertici e archi siano sottoinsiemi di  $\mathcal{V}$  e  $\mathcal{E}$  rispettivamente. Un *sottografo spanning* di  $\mathcal{G}$  è un sottografo che abbia tutti i vertici di  $\mathcal{G}$ .

Inoltre,

**Definizione 1.23.** Sia dato un grafo  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  e siano  $u$  e  $v \in \mathcal{V}$ . Diremo che  $u$  e  $v$  sono *adiacenti* se esiste un arco  $(u, v) \in \mathcal{E}$  tra di loro e l'arco si dirà *incidente* sia  $u$  che  $v$ . Se  $(u, v)$  è un arco orientato, diremo che  $u$  è la sua *origine* e  $v$  la sua *destinazione* ed anche che l'arco *esce* da  $u$  ed *entra* in  $v$ . Se in un arco l'origine e la destinazione coincidono, ad esso ci si riferirà come *cappio*.

Ad esempio, la mappa di una città può essere modellata con un grafo che abbia come vertici gli incroci e i punti in cui finiscono le strade senza uscita e come archi i pezzi di strada -senza intersezioni- che collegano i vertici. Si tratta chiaramente di un grafo misto, in cui si ha archi non diretti (le strade a doppio senso) e diretti (quelle a senso unico).

L'esempio qui riportato, a rigore, non è strettamente aderente alla definizione di grafo appena fornita, in quanto si è caratterizzato  $\mathcal{E}$  come un *insieme*, per cui non si può avere più archi con stessa origine e destinazione, detti *archi paralleli* o *multipli* (ad es., strade parallele che partono ed arrivano agli stessi incroci). La definizione generale, in effetti, prevede che  $\mathcal{E}$  sia una *collezione*, cioè che ammetta anche elementi duplicati al suo interno, mentre i grafi da noi definiti sono detti *semplici*. Per quanto riguarda le reti bayesiane, incontreremo sempre grafi semplici, quindi non ce ne preoccuperemo oltre.

Possiamo a questo punto definire il grado di un vertice.

**Definizione 1.24** (Grado di un vertice). In un grafo non orientato, il *grado* di un vertice è il numero degli archi incidenti sul vertice stesso. Un vertice di grado zero è chiamato *isolato*.

In un grafo orientato, diremo *grado uscente* di un vertice il numero di archi che hanno origine nel vertice e *grado entrante* il numero di archi per cui il vertice è destinazione. Nel caso di un grafo orientato, il suo *grado* sarà la somma del grado uscente e di quello entrante.

Inoltre,

**Definizione 1.25** (Cammino). Viene detto *cammino* da un vertice  $u$  ad un vertice  $u'$  in un grafo  $\mathcal{G}$  una sequenza  $p = \langle v_1, v_2, \dots, v_k \rangle$  (con  $k \geq 2$ ) di vertici tali che  $u = v_1$ ,  $u' = v_k$  e  $(v_{i-1}, v_i) \in \mathcal{E}$  per  $i = 2, \dots, k$ . I nodi  $v_2, \dots, v_{k-1}$  di  $p$  sono detti *interni*. Diremo inoltre che un cammino è *semplice* se tutti i vertici che lo compongono sono distinti.

Se esiste un cammino  $p$  da  $u$  a  $u'$  diremo che  $u'$  è *raggiungibile* da  $u$ .

Un *ciclo* è un cammino, che abbia almeno un arco, in cui  $v_0 = v_k$ . Esso si dirà *semplice* se tutti i vertici, ad eccezione del primo e dell'ultimo, sono distinti.

Un cammino *diretto* è un cammino in cui tutti gli archi sono diretti e sono percorsi lungo la loro direzione. In maniera analoga si definisce un *ciclo diretto*.

Un grafo  $\mathcal{G}$  senza cicli è *aciclico*.

Per indicare la raggiungibilità tra i vertici di un grafo in generale, si introduce la nozione di componenti connesse.

**Definizione 1.26** (Grafo connesso). Un grafo  $\mathcal{G}$  è *connesso* se, per ogni coppia di vertici, c'è un cammino tra di essi. Se  $\mathcal{G}$  non è connesso, i suoi sottografi connessi massimali sono chiamati *componenti connesse di  $\mathcal{G}$* . Esse rappresentano le classi di equivalenza dei vertici secondo la relazione "sono raggiungibili da". Un grafo si dice *fortemente connesso* se ogni vertice è raggiungibile da tutti gli altri nodi. In maniera simile si definiscono le *componenti fortemente connesse*.

**Definizione 1.27** (Versioni orientata e non orientata di un grafo). Sia dato un grafo  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  non orientato. La sua *versione orientata* è il grafo orientato  $\mathcal{G}' = (\mathcal{V}, \mathcal{E}')$  in cui ogni arco non orientato  $(u, v)$  è sostituito dalla coppia di archi orientati  $(u, v)$  e  $(v, u)$ .

Sia dato un grafo  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  orientato. La sua *versione non orientata* è il grafo

non orientato  $\mathcal{G}' = (\mathcal{V}, \mathcal{E}')$  in cui  $(u, v) \in E'_2$  se e solo se  $u \neq v$  e  $(u, v) \in E_2$ .  
 Se un grafo  $\mathcal{G}$  è orientato, diremo *vicino* di un vertice  $u$  un qualsiasi vertice adiacente ad  $u$  nella versione non orientata di  $\mathcal{G}$ .

Diamo alcune definizioni di particolari tipi di grafi.

**Definizione 1.28.** Chiameremo:

- *grafo completo* un grafo non orientato in cui ogni coppia di vertici è adiacente;
- *foresta* un grafo aciclico e non orientato;
- *DAG* un grafo orientato aciclico (da Directed Acyclic Graph).
- *albero (libero)* un grafo connesso, aciclico e non orientato;
- *albero spanning* un sottografo spanning che sia anche un albero libero.

Si noti come la definizione di albero qui introdotta non coincida con la classica nozione di albero (radicato), in cui esiste un nodo, detto appunto *radice*, con particolari proprietà.

**Definizione 1.29** (Relazioni di parentela tra nodi). Dato un DAG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  e due vertici  $u$  e  $v \in \mathcal{V}$ . Diremo che:

- $u$  è *padre* (o *genitore*) di  $v$  se c'è un arco che da  $u$  va in  $v$ . Allo stesso modo,  $v$  viene definito *figlio* di  $u$ ;
- $u$  è detto *antenato* di  $v$  e, viceversa,  $v$  è *discendente* di  $u$ , se esiste un cammino che da  $u$  va in  $v$ . In caso contrario, ci si riferisce a  $v$  come ad un nodo *non discendente*. Si noti che, poichè nella definizione di cammino abbiamo imposto  $k \geq 2$ ,  $u$  non è discendente di se stesso.

Nel caso dei grafi diretti è possibile definire una relazione di ordine.

**Definizione 1.30** (Ordine topologico). Sia dato un grafo diretto  $\mathcal{G}$  con  $n$  vertici. Un *ordine topologico* di  $\mathcal{G}$  è un ordinamento  $v_1, \dots, v_n$  dei suoi vertici tale che, per ogni arco  $(v_i, v_j) \in \mathcal{E}$ ,  $i \leq j$ . In altre parole, si tratta di un ordinamento dei vertici per cui ogni cammino orientato in  $\mathcal{G}$  attraversa i vertici in ordine crescente.

Si noti che un grafo diretto può ammettere più di un ordine topologico.

Vale la seguente proposizione relativa all'ordine topologico dei DAG.

**Proposizione 1.6.** *Un grafo diretto  $\mathcal{G}$  ammette ordine topologico se e solo se è aciclico.*

Infine, ancora un po' di terminologia:

**Definizione 1.31.** Siano dati un DAG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  ed un sottoinsieme di nodi  $\{X_1, X_2, \dots, X_k\} \in \mathcal{V}$ , con  $k \geq 2$ , tali che  $(X_{i-1}, X_i) \in \mathcal{E}$  o  $(X_i, X_{i-1}) \in \mathcal{E}$  per  $2 \leq i \leq k$ . Diremo che l'insieme di archi che collega i  $k$  nodi è una *catena* tra  $X_1$  e  $X_k$  e la indicheremo con entrambe le notazioni  $[X_1, X_2, \dots, X_k]$  e  $[X_k, X_{k-1}, \dots, X_1]$ .

$X_2, \dots, X_{k-1}$  saranno detti *nodi interni*. Una *sottocatena* tra  $X_i$  e  $X_j$  è la catena

$[X_i, X_{i+1}, \dots, X_j]$ , con  $1 \leq i \leq j \leq k$ . Diremo *ciclo* una catena tra un nodo e se stesso, *catena semplice* una catena che non abbia sottocatene che formano cicli. Definiremo *link* una catena che contenga due nodi  $X - Y$ , senza riferimento al verso dell'arco. Un link diretto  $X \rightarrow Y$  verrà indicato come *arco*,  $X$  sarà la sua *coda* e  $Y$  la sua *testa*. Inoltre:

- una catena  $X \rightarrow Z \rightarrow Y$  sarà chiamata *intersezione testa-coda*;
- una catena  $X \leftarrow Z \rightarrow Y$  sarà detta *intersezione coda-coda*;
- una catena  $X \rightarrow Z \leftarrow Y$  sarà chiamata *intersezione testa-testa*;
- una catena  $X - Z - Y$ , con  $X$  e  $Y$  non adiacenti, sarà detta *intersezione non accoppiata*.

## 1.2 Reti bayesiane

Nella precedente sezione abbiamo definito tutte le nozioni di base necessarie per poter introdurre il concetto di rete bayesiana.

Quando non diversamente indicato, i risultati di questa sezione sono adattati da (Pearl 1988) e (Neapolitan 2003), a cui si rimanda per una trattazione più approfondita.

### Definizione

Una condizione fondamentale che deve essere soddisfatta dalle reti bayesiane è la seguente.

**Definizione 1.32** (Condizione di Markov). Siano dati una distribuzione  $P$  delle v.a. in un insieme  $\mathcal{V}$  e un DAG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . Diremo che la coppia  $(\mathcal{G}, P)$  soddisfa la *condizione di Markov* se, per ogni variabile  $X \in \mathcal{V}$ ,  $\{X\}$  è condizionatamente indipendente dall'insieme di tutti i suoi non-discendenti dato l'insieme dei suoi genitori. In maniera equivalente, diremo che  $\mathcal{G}$  e  $P$  soddisfano la condizione di Markov l'uno rispetto all'altro.

In simboli:

$$I_P(\{X\}, ND_X \mid PA_X)$$

dove  $PA_X$  è l'insieme dei genitori (*parents*) di  $X$  e  $ND_X$  è il suo insieme dei non-discendenti.

Si noti che, se  $X$  è una radice (intendendo con ciò che non abbia nodi padre), la condizione di Markov corrisponde semplicemente a dire che è indipendente rispetto ai suoi non-discendenti. Inoltre, se la condizione vale,  $\{X\}$  è condizionatamente indipendente rispetto ad ogni  $S \in ND_X$ . Infine, evidenziamo che i genitori di  $\{X\}$  sono non-discendenti rispetto ad esso, cioè  $PA_X \in ND_X$ . Nel seguito indicheremo un generico padre di  $X$  come  $pa_i \in PA_X$  ed un generico non-discendente come  $nd_i \in ND_X$ .

Se abbiamo una coppia  $(\mathcal{G}, P)$  che sappiamo soddisfare la legge di Markov, possiamo ottenere la distribuzione congiunta  $P$  dalle probabilità condizionate grazie a questo teorema.

**Teorema 1.7.** *Se  $(\mathcal{G}, P)$  soddisfa la condizione di Markov, allora  $P$  è uguale al prodotto delle distribuzioni condizionate di tutti i suoi nodi dati i rispettivi padri, purchè tali distribuzioni esistano. In simboli:*

$$P(x_n, x_{n-1}, \dots, x_1) = P(x_n | pa_n)P(x_{n-1} | pa_{n-1}) \dots P(x_1 | pa_1) \quad (1.23)$$

L'importanza del teorema 1.7 è che, mentre la definizione di distribuzione congiunta è esponenziale nel numero  $n$  di variabili, avendo un DAG in cui ogni nodo abbia al più  $k$  genitori, si passa dalla necessità di dover definire circa  $v^n$  probabilità (in realtà, meno, in base a quante radici ci sono nel DAG e alle reali cardinalità dei nodi) a meno di  $v^k n$ , dove  $v$  indica il massimo numero di valori che i nodi possono assumere. È evidente che, se  $k$  è sufficientemente piccolo, il guadagno è notevole.

Resta però un problema: nella realtà, noi non conosciamo la distribuzione  $P$  nè, men che meno, sappiamo se essa soddisfa la condizione di Markov rispetto ad un certo DAG  $\mathcal{G}$ , in quanto normalmente misuriamo le distribuzioni condizionate. Fortunatamente per noi, il seguente teorema ci viene in soccorso.

**Teorema 1.8.** *Sia dato un DAG  $\mathcal{G}$  in cui ogni nodo sia una v.a. e per ogni suo nodo  $X$  sia fornita una distribuzione di probabilità discreta condizionata rispetto ai valori dei genitori di  $X$  in  $\mathcal{G}$ . In questo caso, avremo che il prodotto di queste probabilità condizionate impone una probabilità congiunta  $P$  sulle variabili e  $(\mathcal{G}, P)$  soddisfa la condizione di Markov.*

Possiamo a questo punto dare la definizione di rete bayesiana.

**Definizione 1.33** (Rete bayesiana). Siano  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  un DAG e  $P$  una distribuzione congiunta di probabilità su  $\mathcal{V}$ . Diremo che  $(\mathcal{G}, P)$  è una *rete bayesiana* se soddisfa la condizione di Markov.

Ci si riferisce a  $\mathcal{G}$  come alla *struttura* della rete ed ai valori assunti dalle distribuzioni condizionate  $P$  come ai suoi *parametri*.

Dai teoremi 1.7 e 1.8 sappiamo che  $P$  è pari al prodotto delle varie distribuzioni condizionate dei nodi e, viceversa, che, dato un DAG e le distribuzioni condizionate discrete (e talvolta continue) dei suoi nodi, possiamo sempre ottenere una distribuzione congiunta  $P$  che soddisfi la condizione di Markov. In seguito, per brevità, ci riferiremo ad una rete bayesiana anche con l'abbreviazione BN (*Bayesian Network*).

Il grande interesse per le reti bayesiane è relativo alla loro capacità -sotto certe condizioni- di determinare legami causali tra le variabili che costituiscono i nodi del grafo. Iniziamo col dare alcune definizioni che ci torneranno utili in seguito.

**Definizione 1.34** (DAG causale). Sia dato un insieme di v.a.  $\mathcal{V}$ . Chiameremo *DAG causale* il grafo che otteniamo se per ogni  $X, Y \in \mathcal{V}$  disegniamo un arco tra  $X$  e  $Y$  se e solo se  $X$  è una causa diretta di  $Y$  relativa a  $\mathcal{V}$ , ossia se  $\nexists W \subseteq \mathcal{V} - \{X, Y\}$  tale per cui, se imponiamo i valori di  $W$ , una manipolazione di  $X$  non modifica più la distribuzione di  $Y$ .

Occorre però formalizzare la definizione di causa.

**Definizione 1.35.** Siano date due v.a.  $X$  e  $Y$ . Definiremo *manipolazione* di  $X$  l'assegnazione ad  $X$  di un determinato valore e diremo che  $X$  *causa*  $Y$  se esiste una manipolazione di  $X$  in grado di modificare la distribuzione di probabilità di  $Y$ .

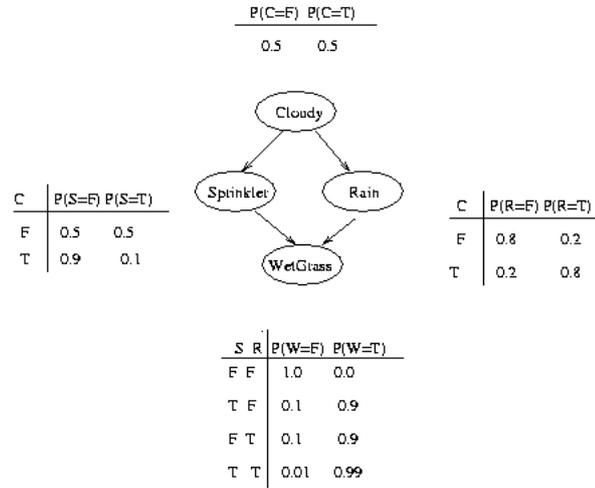


Figura 1.1: Esempio di rete bayesiana (da (Russell e Norvig 2003)).

E' necessario soffermarsi su questo concetto. Stiamo di fatto assumendo che esista una correlazione tra  $X$  ed  $Y$ , ma questo non implica necessariamente che esista un legame di causalità tra le due variabili o che esso sia diretto nel verso  $X \rightarrow Y$ . Infatti, esse risulteranno certamente legate da correlazione sia se  $X$  causa  $Y$ , sia se  $Y$  causa  $X$ , ma anche se esistono entrambe queste relazioni oppure se le due variabili hanno una causa o un effetto in comune. Eseguire una manipolazione significa eseguire un esperimento randomizzato controllato utilizzando una certa popolazione di entità (ad es. sensori rotti) inquadrandoli in un certo contesto (ad es. da quanto sono stati usati e qual è la ditta di fabbricazione). La relazione di causalità avrà senso per *quella popolazione in quel contesto*. La manipolazione rappresenta quindi, in pratica, una relazione di causalità primordiale che è sotto il nostro controllo e per mezzo della quale possiamo scoprire altre relazioni di causalità. Essa non è, tuttavia, il solo modo con cui inferire relazioni di causalità.

Tuttavia, possiamo realmente dire che un DAG causale ed una relativa distribuzione di probabilità formino una rete bayesiana, ovvero soddisfino alla condizione di Markov? In linea teorica, no. Tuttavia, in pratica, è sensato eseguire l'*assunzione della causalità di Markov* se queste condizioni valgono:

1. non ci sono cause comuni nascoste tra i nodi (cioè non esista  $Z \notin V$  tale che  $Z$  causa sia  $X$  che  $Y$ );
2. non c'è bias di selezione;
3. non ci sono cicli di feedback.

Si tenga presente, infine, che la condizione di Markov non è necessariamente legata a relazioni di causalità e perciò può valere anche in assenza di esse. Inoltre, essa implica esclusivamente l'indipendenza, ossia che, se  $(\mathcal{G}, P)$  soddisfa la condizione di Markov e non c'è alcun arco tra  $X$  e  $Y$ , allora certamente non c'è alcuna dipendenza diretta tra le due variabili; la presenza di un arco, tuttavia,

non implica che tale dipendenza esista. Perchè questo valga, dovremo introdurre a breve la nozione di faithfulness.

Tuttavia, la condizione di Markov implica la presenza di altri legami di indipendenza tra le variabili.

**Definizione 1.36** (Indipendenza indotta). Sia data il DAG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . Diremo che, grazie alla condizione di Markov,  $\mathcal{G}$  induce l'indipendenza condizionale  $I_P(A, B | C)$  per  $A, B, C \subseteq V$  se

$$I_P(A, B | C) \quad \forall P \in P_{\mathcal{G}}$$

dove con  $P_{\mathcal{G}}$  indichiamo l'insieme delle distribuzioni per cui  $(\mathcal{G}, P)$  soddisfa la condizione di Markov. Diremo anche che l'indipendenza condizionata è *in*  $\mathcal{G}$ .

Si noti che  $C$  può anche essere l'insieme vuoto, ossia l'indipendenza essere congiunta.

### d-separazione

Per tener conto del legame tra la topologia del grafo e le relazioni di dipendenza che esiste nelle reti bayesiane, si introduce la nozione di d-separazione.

**Definizione 1.37** (Cammini attivi e bloccati). Sia dato un DAG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . Presi tre insiemi  $X, Y, Z \subseteq \mathcal{V}$  diremo che un cammino tra un nodo in  $X$  e un nodo in  $Y$  *attivo* se:

1. ogni nodo con archi convergenti è in  $Z$  o ha un discendente in  $Z$ ;
2. ogni altro nodo è esterno a  $Z$ .

Se almeno una delle due condizioni non è valida, diremo che il cammino è *bloccato*.

**Definizione 1.38** (d-separazione). Sia dato un DAG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . Presi tre insiemi  $X, Y, Z \subseteq \mathcal{V}$  diremo che  $X$  è *d-separato* da  $Y$  tramite  $Z$  se tutti i cammini tra  $X$  e  $Y$  sono bloccati e lo indicheremo con

$$I_{\mathcal{G}}(X, Y | Z)$$

**Definizione 1.39** (Equivalenza di indipendenze condizionali). Siano  $\mathcal{V}$  un insieme di v.a. e  $A_1, B_1, C_1, A_2, B_2, C_2 \subseteq V$ . Diremo che l'indipendenza condizionale  $I_P(A_1, B_1 | C_1)$  è *equivalente* a  $I_P(A_2, B_2 | C_2)$  se per ogni distribuzione  $P$  di  $\mathcal{V}$   $I_P(A_1, B_1 | C_1)$  vale se e solo se vale  $I_P(A_2, B_2 | C_2)$ .

**Definizione 1.40** (Identificazione di d-separazione). Diremo che l'indipendenza condizionale  $I_P(A, B | C)$  è *identificata* da una d-separazione in  $\mathcal{G}$  se vale almeno una delle seguenti condizioni:

- $I_{\mathcal{G}}(A, B | C)$ ;
- $A, B$  e  $C$  non sono mutualmente disgiunti,  $A', B'$  e  $C'$  sono mutualmente disgiunti,  $I_P(A, B | C)$  è equivalente a  $I_P(A', B' | C')$  e  $I_{\mathcal{G}}(A', B' | C')$ .

È possibile dimostrare che la condizione di Markov implica che tutte le d-separazioni sono indipendenze condizionate e, viceversa, che ogni indipendenza condizionata indotta dalla condizione di Markov è identificata da una d-separazione.

### Equivalenza di Markov

A noi interessa sapere anche quando due DAG hanno le stesse d-separazioni e come questo si colleghi alle distribuzioni di probabilità. Introduciamo quindi il concetto di equivalenza di Markov.

**Definizione 1.41** (Equivalenza di Markov). Siano  $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}_1)$  e  $\mathcal{G}_2 = (\mathcal{V}, \mathcal{E}_2)$  due DAG con lo stesso insieme dei vertici. Diremo che  $\mathcal{G}_1$  è *Markov equivalente* a  $\mathcal{G}_2$  se per ogni  $A, B, C \subseteq \mathcal{V}$  disgiunti,  $A$  e  $B$  sono d-separati da  $C$  in  $\mathcal{G}_1$  se e solo se sono d-separati da  $C$  anche in  $\mathcal{G}_2$ , cioè

$$I_{\mathcal{G}_1}(A, B | C) \iff I_{\mathcal{G}_2}(A, B | C)$$

L'importanza di questa definizione per le distribuzioni si deve al seguente teorema.

**Teorema 1.9.** *Due DAG sono Markov equivalenti se e solo se, secondo la condizione di Markov, inducono le stesse indipendenze condizionate.*

Il teorema che caratterizza le equivalenze di Markov è il seguente:

**Teorema 1.10.** *Due DAG  $\mathcal{G}_1$  e  $\mathcal{G}_2$  sono Markov equivalenti se e solo se hanno gli stessi link e lo stesso insieme di intersezioni testa-testa non accoppiate.*

Da questo teorema derivano, da una parte, un algoritmo con tempo polinomiale per testare la Markov-equivalenza tra due DAG (cercando link ed intersezioni testa-testa non accoppiate), dall'altra possiamo rappresentare una classe di equivalenza come un grafo che abbia gli stessi link ed intersezioni testa-testa non accoppiate dei DAG della classe. A quel punto, ogni assegnazione che non crei una nuova intersezione testa-testa non accoppiata o un link diretto è un membro della classe. Diamo le seguenti definizioni:

**Definizione 1.42** (Schema della classe di equivalenza). Dato un DAG  $\mathcal{G}$  chiameremo *schema della classe di equivalenza* (o *pattern DAG*) il grafo che ha gli stessi link dei DAG della classe ed ha orientati solo gli archi comuni a tutta la classe di equivalenza. Diremo che questi archi orientati sono *\*\*\*\*compelled edges\*\*\*\**.

Inoltre,

**Definizione 1.43** (d-separazione per classi di equivalenza). Sia  $gp$  un pattern DAG i cui nodi siano in  $\mathcal{V}$ . Dati tre sottoinsiemi di  $A, B, C \subseteq \mathcal{V}$ , diremo che  $A$  e  $B$  sono *d-separati da  $C$  in  $gp$*  se sono d-separati da  $C$  in almeno uno (e quindi in ogni) DAG della classe di equivalenza definita da  $gp$ .

Per rappresentare le classi di equivalenza si utilizzano i PDAG e i CPDAG:

**Definizione 1.44.** Si definisce *grafo parzialmente diretto aciclico* (PDAG, da Partially Directed Acyclic Graph) un grafo che contenga sia archi diretti che indiretti, ma non cicli diretti. Indicheremo invece con CPDAG un *grafo completo parzialmente diretto aciclico* (da Completed Partially Directed Acyclic Graph) un grafo un PDAG che abbia un arco per ogni arco compelled della classe di equivalenza e un link per ogni arco reversibile. Infine, definiamo *scheletro* di un DAG il grafo indiretto che si ottiene ignorando la direzionalità degli archi.

Si noti che un PDAG non rappresenta necessariamente una classe di equivalenza di DAG, mentre esiste una corrispondenza biunivoca tra CPDAG e classi di equivalenza di DAG.

### Faithfulness

Si era fatto notare precedentemente che la condizione di Markov implica che due nodi  $X$  e  $Y$  di una rete bayesiana tra i quali non ci sia un arco non sono in dipendenza diretta, senza però implicare che la presenza di un arco *garantisca* il legame di dipendenza diretta, come ci aspetteremmo in questo ambito. Dovremo quindi introdurre il concetto di faithfulness, ma prima definiamo formalmente quello di dipendenza diretta.

**Definizione 1.45** (Dipendenza diretta). Data una distribuzione di probabilità  $P$  tra le variabili di un insieme  $\mathcal{V}$  e  $X, Y \in \mathcal{V}$ , diremo che  $X$  e  $Y$  sono in *dipendenza diretta in  $P$*  se  $\nexists W \subseteq V$  per cui  $\{X\}$  e  $\{Y\}$  siano condizionatamente indipendenti.

E, finalmente,

**Definizione 1.46** (Faithfulness). Siano dati un DAG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  ed una distribuzione di probabilità  $P$  sulle variabili in  $\mathcal{V}$ . Diremo che  $(\mathcal{G}, P)$  soddisfa la *condizione di faithfulness* se, secondo la condizione di Markov,  $\mathcal{G}$  induce tutte e sole indipendenze condizionate in  $P$ .

Diremo equivalentemente che  $\mathcal{G}$  è una *mappa perfetta* di  $P$ .

Si ha che

**Teorema 1.11.**  $(\mathcal{G}, P)$  soddisfa la condizione di faithfulness se e solo se tutte le indipendenze condizionate in  $P$  sono identificate da *d-separazione* in  $\mathcal{G}$ .

e

**Teorema 1.12.** Se  $(\mathcal{G}, P)$  soddisfa la condizione di faithfulness, allora  $P$  soddisfa la *faithfulness* per ogni DAG Markov-equivalente a  $\mathcal{G}$ .

### Minimalità

**Definizione 1.47** (Minimalità). Siano dati un DAG  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  ed una distribuzione di probabilità  $P$  su  $\mathcal{V}$ . Diremo che  $(\mathcal{G}, P)$  soddisfa la *condizione di minimalità* se valgono entrambe le condizioni seguenti:

- $(\mathcal{G}, P)$  soddisfa la condizione di Markov;
- rimuovendo un arco da  $\mathcal{G}$ , il DAG risultante non soddisfa più la condizione di Markov su  $P$ .

### Markov Blankets e boundaries

Diamo ora le definizioni di Markov blanket e boundary. Quest'ultimo termine, letteralmente, significa "contorno" ed il concetto che esprime è che, per quanto grande sia la BN in esame, la probabilità di un nodo  $X$  non risente degli effetti di nodi esterni al suo contorno (cioè: è condizionatamente indipendente rispetto ad ogni altro nodo) una volta fissato il valore dei nodi del boundary.

Questa è una definizione leggermente più generale.

**Definizione 1.48** (Markov blanket). Siano dati un insieme  $\mathcal{V}$  di v.a.,  $X \in \mathcal{V}$  e una distribuzione  $P$  su  $\mathcal{V}$ . Chiameremo *Markov blanket*  $M_X$  di  $X$  ogni insieme

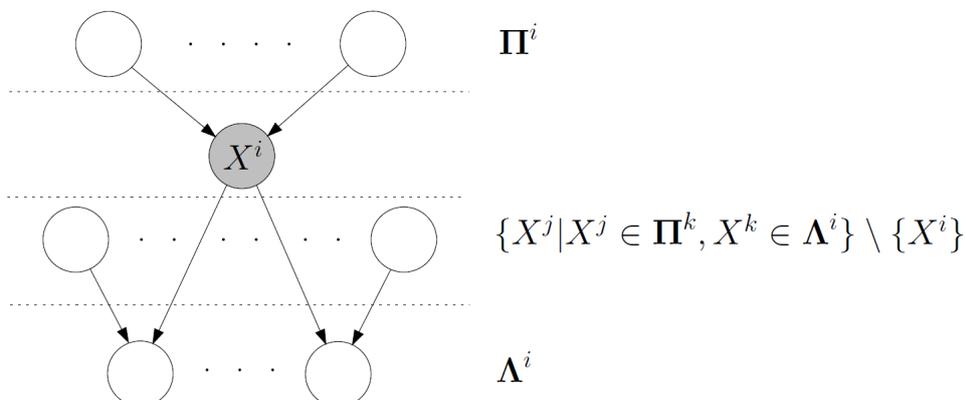


Figura 1.2: Esempio di Markov Blanket. Da (Riggelsen 2006).

di variabili tale che  $X$  sia condizionatamente indipendente da ogni altra variabile dato  $M_X$ , ovvero

$$I_P(\{X\}, \mathcal{V} - (M_X \cup \{X\}) \mid M_X)$$

**Definizione 1.49** (Markov boundary). Siano dati un insieme  $\mathcal{V}$  di v.a.,  $X \in \mathcal{V}$  e una distribuzione  $P$  su  $\mathcal{V}$ . Diremo che una Markov blanket di  $X$  è anche un *Markov boundary* per la variabile se nessuno dei suoi sottoinsiemi propri è Markov blanket per  $X$ .

Segue questo interessante ed utile teorema.

**Teorema 1.13.** *Sia data una coppia  $(\mathcal{G}, P)$  che soddisfi la condizione di Markov. Allora  $\forall X \in \mathcal{V}$ , l'insieme formato dai padri di  $X$ , dai figli di  $X$  e dai padri dei figli di  $X$  è una Markov blanket per  $X$ .*

*Inoltre, se  $(\mathcal{G}, P)$  soddisfa anche la condizione di faithfulness, tale insieme è anche l'unico Markov boundary per  $X$ .*

### Ulteriori cenni

Finora si sono introdotte le principali nozioni necessarie per spiegare cosa sono le reti bayesiane. D'altra parte, in realtà, l'obiettivo per cui sono state sviluppate è per avere uno strumento capace di fornire risposte attendibili ai problemi che presentano la necessità di ragionare probabilisticamente. In altre parole, l'obiettivo è di avere un sistema di *inferenza* che restituisca la distribuzione di probabilità a posteriori relativamente ad un insieme di variabili di interesse dato un ulteriore insieme di variabili osservate. Esistono molteplici metodi per il calcolo dell'inferenza -esatta o approssimata- mediante reti bayesiane, ma per la presente tesi non sono rilevanti e non se ne discuterà.

Inoltre, esistono altre tipologie di reti bayesiane rispetto a quelle trattate in questo testo, quali reti *continue*, in cui i nodi possono assumere un'infinità non numerabile di valori, o reti *dinamiche*, che rappresentano un modello probabilistico temporale (in cui, molto grossolanamente, ogni istante temporale contiene

un modello probabilistico statico). Anche questo non è preso in esame in questa tesi e non se ne parlerà oltre.

Per tutti questi argomenti, un ottimo riferimento può essere il classico (Russell e Norvig 2003).



## Capitolo 2

# Apprendimento con dati incompleti

In questa tesi siamo interessati a valutare la performance di alcuni algoritmi nell'apprendimento della struttura -ossia il DAG- di una rete bayesiana a partire da dati incompleti (o *missing*). Nel presente capitolo introdurremo velocemente il problema generale dell'apprendimento nel caso delle BN, quali sono i processi con cui si può manifestare la mancanza di dati (o *missingness*) e perchè questo crea problemi nell'apprendimento. Seguirà poi la descrizione dei metodi oggetto della comparazione.

### 2.1 Apprendimento di reti bayesiane

Si è detto nel precedente capitolo che uno dei principali obiettivi con cui si introducono le reti bayesiane (e tutte le relative nozioni che riguardano la causalità) è di poter arrivare ad inferire le distribuzioni a posteriori relative a variabili di interesse a partire da dati osservati. Un classico esempio -e un'importante applicazione delle BN nell'ambito medico- è il caso del supporto alla diagnosi. Supponiamo di voler valutare con quale probabilità una persona svilupperà un tumore al polmone. Potremmo quindi considerare come variabili osservabili della nostra BN la professione, l'età, se la persona è fumatrice, ecc., e come variabile incognita la probabilità di sviluppare o meno tale malattia. Assunto che questo sia l'insieme  $\mathcal{V}$  delle variabili del nostro DAG, come possiamo unire i nodi secondo legami di causalità e come possiamo arrivare a conoscere la distribuzione  $P$  della rete in modo da poter inferire la probabilità a posteriori per un determinato soggetto? In un primo caso, potremmo chiedere ad un oncologo di basarsi sulla sua esperienza e tracciare l'intera rete bayesiana (struttura e parametri). In alternativa, potremmo chiedergli di valutare esclusivamente i legami di causalità tra i nodi. In quest'ultimocaso, rimarrebbe solo il problema di imparare la distribuzione  $P$ , ovvero quelli che sono detti *parametri* della BN. Per far ciò, raccoglieremo un campione di individui per cui tutte le variabili -compreso l'aver o meno sviluppato un tumore- siano note, ovvero abbiamo bisogno di un *training set*. A questo punto, utilizzeremo un algoritmo per l'*apprendimento di parametri*.

D'altra parte, non sempre si ha a disposizione un esperto oppure l'insieme di variabili che si sta studiando potrebbe essere così ampio che neanche un esperto potrebbe riuscire a collegare causalmente tutti i nodi. Inoltre, in ogni caso, i legami che si ottengono da un esperto sono comunque frutto di valutazioni soggettive e potenzialmente in disaccordo con i dati che si sta esaminando. In questo caso occorre un *apprendimento di strutture* (o *strutturale*).

### 2.1.1 Il problema della missingness

Nell'introdurre la sezione abbiamo detto che necessitiamo di un training set, che useremo per imparare struttura e/o parametri della BN, in cui tutti i valori siano noti. Com'è però facile immaginare, nella realtà le cose non funzionano in questo modo ed occorrerà trovare il modo di apprendere una BN anche a partire da dati incompleti, come vedremo nella sezione successiva. Ma come e perchè potremmo avere dati mancanti? Ritorniamo all'esempio precedente e supponiamo che, nel rilevare i dati relativi ai soggetti del nostro training set, un'infermiera distratta ogni tanto si dimentichi di chiedere l'età ad un paziente, se è fumatore ad un altro e così via. In questo caso, i dati mancheranno in modo del tutto casuale ed imprevedibile. Potremmo, d'altra parte, rilevare che i pazienti "sani" che sviluppano normali bronchiti tendono a non avere lastre toraciche, in quanto il loro medico non scorge pericolo che queste nascondano un tumore e non prescrive l'esame. In questo caso, il valore assunto da una variabile (l'intensità della bronchite) influisce sull'assenza di un'altra (le lastre). Infine, potremmo avere che i soggetti che presentano disfunzioni sessuali per pudore tendano ad omettere le risposte relative all'argomento. In questo caso, il valore assunto dalla variabile influisce sull'assenza di dati della stessa. Per uno studio sul valore informativo della missingness in ambito medico, si può vedere (Lin e Haug 2008).

L'esempio appena riportato illustra i tre tipi di missingness indicati da (Little e Rubin 1986), che ora formalizziamo. Iniziamo indicando con  $\mathcal{D}$  una matrice che rappresenta l'insieme dei dati di ingresso, di dimensioni  $d \times N$ , con  $d$  pari al numero di vettori di dati ed  $N$  la dimensione di ciascun vettore, e con  $\mathcal{M}$  una matrice binaria  $d \times N$ , detta *matrice indicatrice dei dati mancanti*, in cui l'elemento  $m_{ij}$  vale uno se il dato  $d_{ij}$  è assente. Tale matrice suddivide  $\mathcal{D}$  in due parti:  $\mathcal{D}_o$ , cioè l'insieme dei vettori che hanno tutti i dati, e  $\mathcal{D}_m$ , cioè l'insieme dei vettori con alcuni dati mancanti. In generale, il meccanismo per la generazione di dati mancanti è rappresentabile come una distribuzione condizionata di  $\mathcal{D}$  dato  $\mathcal{M}$  e l'insieme  $\xi$  dei parametri incogniti che definisce il meccanismo di generazione della missingness:

$$p(\mathcal{M} \mid \mathcal{D}, \xi) = p(\mathcal{M} \mid \mathcal{D}_o, \mathcal{D}_m, \xi)$$

Possiamo ora dare la definizione formale:

**Missing Completely At Random (MCAR).** Quando la probabilità di assenza è indipendente sia dai valori assunti dalla variabile stessa, sia da quelli assunti da ogni altra. In questo modo, i dati disponibili sono sufficienti ad eseguire le inferenze, in quanto la missingness è indipendente dai dati in ingresso. In formule:

$$p(\mathcal{M} \mid \mathcal{D}_o, \mathcal{D}_m, \xi) = p(\mathcal{M} \mid \xi)$$

**Missing At Random (MAR).** L'assenza del dato è indipendente dal valore assunto dalla variabile mancante, ma risulta esserci una qualche forma di dipendenza dai dati osservati, ovvero:

$$p(\mathcal{M} \mid \mathcal{D}_o, \mathcal{D}_m, \xi) = p(\mathcal{M} \mid \mathcal{D}_o, \xi)$$

**Not Missing At Random (NMAR).** Anche in questo caso esiste uno schema non casuale secondo cui i dati sono mancanti ma, a differenza del caso MAR, la missingness dipende dal valore assunto dalla variabile mancante stessa e quindi l'informazione non può essere ricostruita basandosi sui soli dati osservati.

Nei casi MCAR e MAR si è soliti dire che il meccanismo di incompletezza è *ignorabile*. Si tratta, in linea generale, di casi più favorevoli rispetto ad un'incompletezza di tipo NMAR ed infatti la maggior parte dei metodi tende ad assumere che il meccanismo di missingness sia ignorabile.

### 2.1.2 Apprendimento di parametri

In questo paragrafo assumiamo nota la struttura della nostra rete bayesiana, ad esempio perchè le relazioni tra le variabili sono state fissate da un esperto del settore a cui la rete si riferisce. Il nostro obiettivo sarà quindi, noto il DAG  $\mathcal{G}$ , trovare i parametri della rete bayesiana  $(\mathcal{G}, P)$ , ossia determinare la densità congiunta  $P$  (o, equivalentemente e più realisticamente, tutte le distribuzioni condizionate necessarie per ricostruirla). I risultati qui riportati sono limitati all'essenziale ed adattati da (Neapolitan 2003) e (Russell e Norvig 2003), a cui si rimanda per una trattazione decisamente più approfondita.

#### Approccio bayesiano

L'apprendimento dei parametri -ed in larga parte anche della struttura- di un rete bayesiana utilizza un approccio bayesiano. In parole povere, questo significa che viene calcolata la probabilità di ogni ipotesi condizionatamente ai dati e le predizioni vengono fatte su tale base, usando cioè tutti le ipotesi pesate per le loro probabilità (e non solo una singola ipotesi migliore). In pratica, il problema dell'apprendimento viene ricondotto ad un problema di inferenza, come si è visto nel capitolo 1 (1.1.1). Se abbiamo un dataset  $\mathcal{D}$ , di cui indicheremo i valori osservati con  $\mathbf{d}$ , dalla legge di Bayes avremo

$$\begin{aligned} P(h_i \mid \mathbf{d}) &= \frac{P(\mathbf{d} \mid h_i)P(h_i)}{P(\mathbf{d})} \\ &= cP(\mathbf{d} \mid h_i)P(h_i) \end{aligned} \quad (2.1)$$

dove con  $h_i$  abbiamo indicato l'ipotesi  $i$ -esima e con  $c$  il termine  $\frac{1}{P(\mathbf{d})}$ , in quanto non dipende dalle ipotesi ed è perciò comune a tutti i casi (cioè non interessante al fine di discernere l'ipotesi migliore). Indicando con  $X$  una quantità non nota, dal teorema della probabilità totale si ha inoltre

$$P(X \mid d) = \sum_i P(X \mid d, h_i)P(h_i \mid d) \quad (2.2)$$

Come si può vedere, le ipotesi servono a modulare le predizioni rispetto ai dati grezzi e le quantità che di esse ci interessano sono:

- **Prior**, cioè la probabilità a priori  $P(h_i)$  che si verifichi l'ipotesi  $h_i$ ;
- **Likelihood** (o *verosimiglianza*), cioè la probabilità  $P(\mathbf{d} | \mathbf{h}_i)$  di ottenere i dati che si sta analizzando sapendo che si è verificata l'ipotesi  $i$ -esima;
- **Posterior**, che si ricava da prior e verosimiglianza tramite la legge di Bayes, come indicato nell'equazione 2.1.

Facciamo inoltre un'altra ipotesi, cioè che le osservazioni siano *indipendenti ed identicamente distribuite (i.i.d.)*, di modo che valga

$$P(\mathbf{d} | h_i) = \prod_j P(d_j | h_i)$$

Si noti che, in presenza di molti dati (al limite, infiniti), anche un eventuale prior sbagliato influisce poco sulla predizione, tendendo verso l'ipotesi suggerita dai dati. In caso di pochi dati, invece, se il prior è corretto la predizione può migliorare sensibilmente rispetto ad approcci deterministici (a scapito di una maggior richiesta computazionale). C'è però un problema: nei casi reali lo spazio delle ipotesi può essere molto grande se non infinito, rendendo di fatto improponibile la somma 2.2 (o il corrispondente integrale, nel caso continuo). Si deve dunque eseguire un'approssimazione, che spesso è nella forma *Maximum A Posteriori (MAP)*. In questo caso, la predizione si basa su una sola ipotesi  $h_{MAP}$  che abbia il massimo valore della posterior  $P(h_i | \mathbf{d})$  rispetto a tutte le altre ipotesi. Si ha perciò un'approssimazione della stima bayesiana in cui  $P(X | \mathbf{d}) \approx P(X | h_{MAP})$ . Un'ulteriore approssimazione si può avere se si può assumere un prior uniforme (in cui, cioè, nessuna ipotesi sia a priori più probabile di un'altra), per cui  $P(h_i) = \frac{1}{N_h} \quad \forall \quad i = 1, \dots, N_h$  (con  $N_h$  numero delle ipotesi) può non essere preso in considerazione nel massimizzare l'ipotesi MAP. Quando ciò avviene, si sta eseguendo un'approssimazione *Maximum Likelihood (ML)* (o *Massima Verosimiglianza*), in cui si cerca una ipotesi  $h_{ML}$  che massimizzi  $P(\mathbf{d} | h_i)$ , com'è evidente dalla legge di Bayes. Allo stesso modo, si può cercare di massimizzare la log-likelihood, che semplifica i conti trasformando prodotti in somme:

$$L(\mathbf{d} | h_i) = \log P(\mathbf{d} | h_i) = \sum_j \log P(d_j | h_i) \quad (2.3)$$

### Funzioni beta e gamma

Introduciamo qui alcune funzioni che ci verranno utili nel paragrafo successivo. Ci limiteremo al caso più semplice di variabili binarie, onde evitare di appesantire notazioni e concetti. L'estensione al caso di variabili in grado di assumere più di due valori è immediato e si rimanda il lettore alla bibliografia.

**Definizione 2.1** (Funzione gamma). La *funzione gamma*  $\Gamma : \mathbb{R} \rightarrow \mathbb{R}$  è definita come:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad (2.4)$$

L'integrale converge se e solo se  $x \geq 0$ . Inoltre

$$\frac{\Gamma(x+1)}{\Gamma(x)} = x$$

e, se  $x \geq 1$  e intero, si ha

$$\Gamma(x) = (x - 1)!$$

**Definizione 2.2** (Funzione beta). La funzione  $\rho : \mathbb{R} \rightarrow \mathbb{R}$  con parametri  $a$  e  $b \in \mathbb{R}$ ,  $a, b \geq 0$  e  $N = a + b$  è definita come:

$$\rho(f) = \frac{\Gamma(N)}{\Gamma(a)\Gamma(b)} f^{a-1} (1-f)^{b-1}, \quad 0 \leq f \leq 1 \quad (2.5)$$

Si dice che una v.a.  $F$  che ha una densità di probabilità  $\rho$  ha una *distribuzione beta*. Faremo riferimento a  $a$  e  $b$  anche come ad *iperparametri*, in quanto sono parametri della distribuzione relativa ai parametri della BN.

Il valore atteso di una distribuzione beta è

$$E[F] = \frac{a}{N}$$

Se abbiamo una v.a.  $X$  che può assumere valori 1 e 2, vale il seguente teorema.

**Teorema 2.1.** *Se  $X$  è una v.a. che può assumere valori 1 e 2 ed  $F$  è una v.a. con distribuzione beta di parametri  $a$  e  $b$  come sopra, allora se*

$$P(X = 1 | f) = f$$

si ha

$$P(X = 1) = E[F] = \frac{a}{N}$$

Diamo la definizione di campione binomiale.

**Definizione 2.3** (Campione binomiale). Chiameremo *campione binomiale di taglia  $M$  e parametro  $F$*  un insieme di variabili (o vettori) aleatori  $\mathbf{D} = \{X_1, X_2, \dots, X_M\}$ , tutti con lo stesso spazio  $\{1, 2\}$  se esiste una v.a.  $F$  con spazio  $[0, 1]$  e funzione di densità  $\rho$  tale che tutti gli  $X_i$  sono a due a due condizionatamente indipendenti rispetto a  $F$ , tale che

$$P(X_i = 1 | f) = f \quad \forall \quad 1 \leq i \leq M$$

Ora possiamo arrivare al seguente teorema, estremamente importante per l'apprendimento parametrico.

**Teorema 2.2.** *Siano:*

1.  $\mathbf{D}$  un campione binomiale di taglia  $M$  e parametro  $F$ ;
2.  $\mathcal{D}$  un insieme di valori in  $\mathbf{D}$  che chiameremo dataset

$$\mathcal{D} = \{x_1, \dots, x_M\}$$

3.  $s$  il numero di variabili in  $\mathcal{D}$  pari ad 1;
4.  $t$  il numero di variabili in  $\mathcal{D}$  pari a 2;
5.  $F$  con distribuzione beta di parametri  $a$  e  $b$ ,  $N = a + b$ .

Allora:

$$P(\mathcal{D}) = \frac{\Gamma(N)}{\Gamma(N+M)} \frac{\Gamma(a+s)\Gamma(b+t)}{\Gamma(a)\Gamma(b)} \quad (2.6)$$

e

$$\rho(f | \mathcal{D}) = \text{beta}(f; a+s, b+t) \quad (2.7)$$

L'equazione 2.7 ci dice che, se aggiorniamo una funzione di densità beta relativa ad un campione binomiale tenendo conto dei dati in esso contenuti, otteniamo un'altra funzione di densità beta. Si usa quindi dire che l'insieme delle funzioni beta è una *famiglia coniugata di funzioni di densità*.

Ed infine, si fornisce qui una generalizzazione della funzione beta:

**Definizione 2.4** (Funzione di densità di Dirichlet). La *funzione di densità di Dirichlet* di parametri  $a_1, a_2, \dots, a_r \in \mathbb{N}$  e  $a_1, a_2, \dots, a_r \geq 1$  con  $N = \sum_{k=1}^r a_k$  è definita come

$$\rho(f_1, f_2, \dots, f_{r-1}) = \frac{\Gamma(N)}{\prod_{k=1}^r \Gamma(a_k)} f_1^{a_1-1} f_2^{a_2-1} \dots f_r^{a_r-1} \quad 0 \leq f_k \leq 1, \sum_{k=1}^r f_k = 1 \quad (2.8)$$

La funzione di densità di Dirichlet è indicata  $\text{Dir}(f_1, f_2, \dots, f_{r-1}; a_1, a_2, \dots, a_r)$  (o, per semplicità, anche solo  $\text{Dir}(a_1, a_2, \dots, a_r)$ ) e si dice che le v.a.  $F_1, F_2, \dots, F_r$  con questa densità di probabilità hanno *distribuzione di Dirichlet*.

Si noti che, poichè la somma delle  $f_i$  è pari ad uno, il valore di  $f_r$  (e perciò di  $F_r$ ) è univocamente determinato dalle prime  $r-1$  variabili.

In questo caso, ad essere indicati come iperparametri saranno gli  $a_1, a_2, \dots, a_r$ . La funzione beta non è altro che una Dirichlet in cui  $r=2$ .

E' importante far presente che, come nel caso della funzione beta, la distribuzione di Dirichlet aggiornata utilizzando i dati, che indicheremo con  $\rho | \mathcal{D}$ , resta una Dirichlet e si modifica in modo simile alla densità beta. Cioè, se  $\rho = \text{Dir}(f_1, \dots, f_{r-1}; a_1, \dots, a_r)$  e  $s_1, \dots, s_r$  indicano il numero di volte  $s_i$  in cui  $a_i$  appare nel dataset (definite anche *statistiche sufficienti* per la distribuzione di Dirichlet), allora

$$\rho | \mathcal{D} = \text{Dir}(f_1, \dots, f_{r-1}; a_1 + s_1, a_2 + s_2, \dots, a_r + s_r)$$

### Caso 1: dati completi

I metodi che presenteremo suppongono che i parametri siano distribuiti secondo leggi di Dirichlet, pertanto qui indichiamo soltanto come ottenere le distribuzioni condizionate in tal caso. Altri casi sono reperibili in letteratura ma non verranno qui presi in considerazione.

Consideriamo una rete bayesiana  $(\mathcal{G}, P)$  con  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  noto. Questa è la simbologia che ci servirà:

- $N$  sarà la cardinalità di  $\mathcal{V}$  e  $V_i$  sarà un elemento di tale insieme;
- $r_i$  sarà il numero di valori che potrà assumere il vertice  $V_i$ ;
- $pa_{ij}$  sarà la  $j$ -esima configurazione dei padri del nodo  $V_i$ . Ad esempio, se  $V_i$  ha due genitori ciascuno dei quali può assumere due valori, l'insieme delle configurazioni sarà  $\{pa_{ij}\} = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$  e diremo quindi che, ad es.,  $pa_{i3} = (2, 1)$ ;

- $q_i$  sarà il numero di configurazioni che possono assumere i genitori di  $V_i$ . Nell'esempio precedente,  $q_i = 4$ ;
- $p_{ijk}$  indica la probabilità assunta dal nodo  $V_i$  quando i suoi genitori sono in configurazione  $pa_{ij}$  e  $V_i = k$ , con  $k = 1, 2, \dots, r_i$ .  $p_{ijk}$  è quindi una voce della CPT relativa al nodo  $V_i$ ;
- $a_{ijk}$  indicherà l'iperparametro della distribuzione di Dirichlet relativa al nodo  $V_i = k$  con configurazione dei padri pari a  $pa_{ij}$ .  $a_{ij} = \sum_{k=1}^{r_i} a_{ijk}$ ;
- $s_{ijk}$  sono le statistiche sufficienti per la distribuzione di Dirichlet e  $N_{ij} = \sum_{k=1}^{r_i} s_{ijk}$ ;
- $\mathcal{D}$  rappresenta il nostro dataset.

A questo punto, per ogni nodo, definiremo i valori a priori della relativa distribuzione di Dirichlet, ossia i vari  $a_{ijk}$ . Poichè  $\mathcal{G}$  è dato, le probabilità a priori saranno:

$$P(p_{ij1}, \dots, p_{ijr_i} \mid \mathcal{G}) = \text{Dir}(a_{ij1}, \dots, a_{ijr_i}) = \Gamma(a_{ij}) \prod_{k=1}^{r_i} \frac{p_{ijk}^{a_{ijk}-1}}{\Gamma(a_{ijk})}$$

Assumendo l'*indipendenza globale e locale* dei parametri (Spiegelhalter e Lauritzen 1990), (Cooper e Herskovits 1992), (Heckerman, Geiger e Chickering 1995), sull'insieme  $\mathbf{p}$  dei parametri della BN si ha

$$P(\mathbf{p} \mid \mathcal{G}) = \prod_{i=1}^N \prod_{j=1}^{q_i} \Gamma(a_{ij}) \prod_{k=1}^{r_i} \frac{p_{ijk}^{a_{ijk}-1}}{\Gamma(a_{ijk})}$$

Condizionando poi sul dataset  $\mathcal{D}$ , si ottiene un'ulteriore funzione di Dirichlet, in quanto famiglia coniugata di distribuzioni:

$$P(p_{ij1}, \dots, p_{ijr_i} \mid \mathcal{G}, \mathcal{D}) = \text{Dir}(a_{ij1} + s_{ij1}, \dots, a_{ijr_i} + s_{ijr_i})$$

e

$$P(\mathbf{p} \mid \mathcal{G}, \mathcal{D}) = \prod_{i=1}^N \prod_{j=1}^{q_i} \Gamma(a_{ij} + N_{ij}) \prod_{k=1}^{r_i} \frac{p_{ijk}^{s_{ijk} + a_{ijk} - 1}}{\Gamma(a_{ijk} + s_{ijk})}$$

Spesso, per semplicità, invece dell'intera distribuzione si utilizzano direttamente i valori MAP, per cui la stima è

$$p_{ijk} = \frac{a_{ijk} + s_{ijk}}{a_{ij} + N_{ij}} \quad (2.9)$$

che si semplifica ulteriormente nel caso di grandi campioni ( $N_{ijk} \gg a_{ijk}$ ) o di stima ML:

$$p_{ijk} = \frac{s_{ijk}}{N_{ij}} \quad (2.10)$$

Quanto appena descritto può essere formalmente rappresentato come una rete bayesiana (detta *aumentata*) che incorpora la rete in esame, aggiungendovi per ogni  $a_{ij}$  un vertice radice come genitore del nodo  $V_i$ , tale che ogni vertice aggiunto abbia un solo arco uscente, diretto verso  $V_i$ , che ne indica la distribuzione a priori. Inoltre, nascosto nelle pieghe di queste formule, c'è il modo per definire

quanto forte è il nostro prior. Infatti, quanto più siamo confidenti che un valore  $k$  di  $V_i$  si presenti quando vale  $pa_{ij}$ , tanto più possiamo aumentare il valore del relativo  $a_{ijk}$ . In altre parole, possiamo considerare  $a_{ijk}$  come un *contatore virtuale* che rappresenta -idealmente- quante volte “abbiamo visto” (a priori) che tale caso si è verificato. C'è però un problema. Consideriamo, per semplicità, di voler definire un prior per un nodo binario come una funzione beta. Volendo dare un prior uniforme, potrebbe sembrare indifferente definire  $\text{beta}(f; 1,1)$  o  $\text{beta}(f; 1000,1000)$ , in quanto in entrambi i casi le probabilità a priori per il caso 1 (rappresentato da  $a$ ) e il caso 2 (rappresentato da  $b$ ) saranno 0.5. Però, per quanto appena detto,  $a$  e  $b$  sono dei contatori virtuali e ciò implica che, nell'aggiornare la distribuzione sulla base dei dati, il primo prior è più debole del secondo. Quando si ha a che fare con più nodi, magari inizializzati con valori di  $a_{ij}$  diversi, questo può condurre a risultati spiacevoli. Per evitare ciò, in determinate circostanze si può introdurre la nozione di *ESS* (Equivalent Sample Size, cioè taglia del campione equivalente), per cui gli  $a_{ijk}$  di ogni nodo sono determinati in modo che la loro somma restituisca lo stesso  $a_{ij}$  per ogni vertice della rete.

Le molte nozioni formali e tanti altri risultati teorici per arrivare a tale procedimento (reti bayesiane aumentate e embedded, indipendenza locale e globale dei parametri, ESS, ...) si trovano ad es. in (Neapolitan 2003) ed in parte in (Heckerman, Geiger e Chickering 1995).

### Caso 2: dati incompleti

Nel caso di dati incompleti, la strada solitamente seguita consiste nel ricondursi al caso di dati completi sopperendo alla mancanza di un caso mediante valutazioni statistiche sui casi restanti. Come si è visto in precedenza, però, esistono varie tipologie di missingness e non tutte consentono valutazioni basate esclusivamente sui dati osservabili. Si illustra qui soltanto il metodo *EM* (sigla per *Expectation Maximization*) (Dempster, Laird e Rubin 1977), che consiste nell'iterazione di due passi ed assume che il meccanismo con cui mancano i dati sia di un tipo ignorabile. Per altri metodi, quali Markov Chain Monte Carlo (MCMC) e Gibbs Sampling (Geman e Geman 1984), Data Augmentation (Tanner e Wong 1987) e Bound and Collapse (Ramoni e Sebastiani 1997), si rimanda alla letteratura.

L'intuizione dietro all' algoritmo EM è semplice: non avendo a disposizione tutti i dati, si sfruttano le conoscenze statistiche a disposizione per ricondursi al caso di dati completi, quindi si cercano i parametri in modo da massimizzare la verosimiglianza dei dati. Questi due passaggi saranno ripetuti fino ad arrivare a convergenza (cioè fino a trovare i parametri che massimizzano globalmente la verosimiglianza). Vediamo la cosa un po' più in dettaglio.

Supponiamo noto  $\mathcal{G}$  ed un dataset  $\mathcal{D} = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(M)}\}$ , in cui ogni vettore  $\mathbf{x}^{(h)}$  ha per elementi le v.a.  $X_i$  con  $i = 1, \dots, r_i$ . Inizializziamo il vettore dei parametri  $\mathbf{p}^0$  (il pedice indica il valore del vettore dei parametri all' iterazione  $t$ -esima) imponendo prior di tipo Dirichlet. Poichè la procedura potrebbe fermarsi in un minimo locale, non abbiamo a priori un modo per definire  $\mathbf{p}^0$ . A questo punto, per ogni  $t$  fino a convergenza (garantita sotto determinate condizioni; in pratica molto spesso):

**Expectation** sfruttando i parametri  $\mathbf{p}^t$  dell' iterazione  $t$ -esima, si calcola l'a-

spettazione (da cui il nome) delle statistiche sufficienti  $s_{ijk}$  come

$$s_{ijk}^t = E[s_{ijk} \mid \mathcal{D}, \mathbf{p}^t] = \sum_{h=1}^M P(X_i^{(h)} = k, pa_{ij} \mid \mathbf{x}^{(h)}, \mathbf{p}^t)$$

**Maximization** si utilizzano le statistiche sufficienti  $s_{ijk}^t$  appena trovate e si calcola la stima ML dei parametri  $\mathbf{p}^{t+1}$ , ossia si risolve secondo  $\mathbf{p}$  il sistema

$$E[s_{ijk}^t(\mathcal{D}) \mid \mathbf{p}] = s_{ijk}^t$$

che significa, come si è visto nell'equazione 2.10,

$$p_{ijk} = \frac{s_{ijk}}{N_{ij}}$$

Per casi più generali, si faccia riferimento all'articolo originale (Dempster, Laird e Rubin 1977).

### 2.1.3 Apprendimento di strutture

Nel caso in cui ad essere incogniti non siano soltanto i parametri ma anche la struttura della rete, abbiamo a disposizione metodiche basate su due diversi approcci. Il primo è definito *score based* e si basa sul cercare un insieme di struttura e parametri che massimizzi una funzione punteggio (score, appunto), l'altro è definito *constraint based* e cerca la struttura imponendo vincoli (*constraints*) di indipendenza tra le variabili. Anche in questo caso, il problema della missingness dei dati cambia il modo in cui si può cercare una soluzione, principalmente perchè i metodi basati sui vincoli non possono essere usati e le funzioni punteggio per il caso di dati completi sono scelte per essere decomponibili, in modo che una modifica locale non richieda il ricalcolo del punteggio su tutta la rete, ma questa proprietà viene meno in caso di dati mancanti. Non è il caso di soffermarsi in dettaglio sui metodi, in quanto quelli che ci serviranno saranno oggetto di spiegazione più approfondita nella sezione successiva. Qui ci si limita pertanto a descrivere soltanto le caratteristiche generali del problema.

#### Approccio score based

L'apprendimento di struttura qui in esame può spesso essere visto come un problema di *selezione di modello* in cui, note le variabili in gioco ed avendo a disposizione un dataset  $\mathcal{D}$ , si vuole determinare e selezionare il/i DAG pattern con maggior probabilità a posteriori condizionatamente ai dati, così da poterlo usare a scopo di inferenza. Non è però il solo approccio possibile, in quanto si potrebbe essere interessati ad un punto di vista bayesiano, come in (Riggelsen e Feelders 2005). Inoltre, non è detto che esista un pattern DAG evidentemente migliore degli altri e, nel qual caso può essere più saggio eseguire una media sulle varie strutture trovate. Tuttavia, per il momento non ci occuperemo di questi dettagli.

Definiremo *score* per un DAG  $\mathcal{G}$  o un pattern DAG *gp* una funzione che assegna un valore ad ogni struttura presa in considerazione in base ai dati disponibili.

$$score(\mathcal{G}, \mathcal{D}) = P(\mathcal{G} \mid \mathcal{D}) \quad (2.11)$$

Per la legge di Bayes:

$$P(\mathcal{G} | \mathcal{D}) = \frac{P(\mathcal{D} | \mathcal{G})P(\mathcal{G})}{P(\mathcal{D})}$$

Poichè  $\mathcal{D}$  non dipende dalla struttura, la massimizzazione dello score è pari alla massimizzazione del solo numeratore. In questa sede supponiamo inoltre di avere un prior uniforme sulle strutture, per cui a noi interessa massimizzare soltanto  $P(\mathcal{D} | \mathcal{G})$ . Chiameremo perciò *criterio di score bayesiano* (Bayesian Dirichlet equivalent with uniform priors) BDeu (Heckerman, Geiger e Chickering 1995) il valore

$$\begin{aligned} score_B(\mathcal{D}, gp) &= P(\mathcal{D} | gp) = score_B(\mathcal{D}, \mathcal{G}) = P(\mathcal{D} | \mathcal{G}) \\ &= \prod_{i=1}^n \prod_{j=1}^{q_i^{(\mathcal{G})}} \frac{\Gamma(N_{ij}^{(\mathcal{G})})}{\Gamma(N_{ij}^{(\mathcal{G})} + \Gamma(M_{ij}^{(\mathcal{G})}))} \prod_{k=1}^{r_i} \frac{\Gamma(a_{ijk}^{(\mathcal{G})} + s_{ijk}^{(\mathcal{G})})}{\Gamma(a_{ijk}^{(\mathcal{G})})} \end{aligned} \quad (2.12)$$

Un'importante caratteristica di questo score è che esso è decomponibile, ovvero può essere ottenuto come prodotto di più fattori locali di score:

$$score_B(\mathcal{D}, gp) = score_B(\mathcal{D}, \mathcal{G}) = \prod_{i=1}^n score_B(d, X_i, PA_i^{(\mathcal{G})}) \quad (2.13)$$

dove

$$score_B(d, X_i, PA_i) = \prod_{j=1}^{q^{(PA_i)}} \frac{\Gamma(\sum_k a_{ijk}^{(PA_i)})}{\Gamma(\sum_k a_{ijk}^{(PA_i)} + \sum_k s_{ijk}^{(PA_i)})} \prod_{k=1}^{r_i} \frac{\Gamma(a_{ijk}^{(PA_i)} + s_{ijk}^{(PA_i)})}{\Gamma(a_{ijk}^{(PA_i)})} \quad (2.14)$$

in cui con  $PA_i$  si è indicato l'insieme dei genitori del nodo  $X_i$ , con  $q^{PA_i}$  il numero di possibili combinazioni di valori dei genitori e le altre quantità in modo simile a quanto già visto. In particolare, nel caso del BDeu, in cui i prior sono uniformi, i conteggi immaginari sono pari a  $\frac{ESS}{q_i^{PA_i} r_i}$ , in cui  $r_i$  indica la cardinalità del nodo.

Nel caso di dataset sufficientemente grandi, è possibile eseguire alcune approssimazioni tramite una gaussiana multivariata su questo score, fino ad arrivare al *Bayesian Information Criterion (BIC)* (Schwarz 1978), che può essere a sua volta usato come score:

$$BIC_{score}(\mathcal{G}, \mathcal{D}) = \log P(\mathcal{D} | \hat{\mathbf{p}}, \mathcal{G}) - \frac{d}{2} \log N \quad (2.15)$$

in cui  $\hat{\mathbf{p}}$  rappresenta la stima ML dei parametri della rete e  $d = \sum_{i=1}^N q_i(r_i - 1)$  indica il numero di parametri liberi della gaussiana multivariata. Questo score ha il vantaggio di non dipendere dai prior dei parametri e di avere l'interpretazione intuitiva di verosimiglianza meno un termine di penalità, così da prevenire overfitting e scoraggiare strutture troppo complicate. Un ulteriore score è il *Minimum Description Length (MDL)* (Rissanen 1987) che si basa su concetti di compressione dati ma che è risultato essere equivalente all'opposto del BIC:  $MDL = -BIC$ . Tuttavia, vanno fatte notare alcune osservazioni in (friedman:1998) riguardo a come lo score bayesiano fornisca migliori valutazioni delle proprietà di generalizzazione di un modello condizionatamente ai dati e che esso consenta l'incorporazione di conoscenza a priori nel processo di apprendimento.

Esistono poi ulteriori punteggi, quali AIC (Akaike Information Criterion) (Akaike 1973) e Cheeseman-Stutz Score (Cheeseman e Stutz 1995), che qui non prenderemo in considerazione.

A questo punto, il nostro obiettivo sarà quello di muoverci nell'insieme dei possibili DAG fino a trovare quello con lo score migliore. Una possibile soluzione ingenua è ovviamente quella di verificare una per una tutte le possibili strutture ottenibili a partire dall'insieme dei vertici  $\mathcal{V}$ , ma tale approccio a forza bruta non è praticabile se non con un numero molto piccolo di variabili. In maniera più intelligente, potremmo limitarci all'insieme dei pattern DAG, visto che tutti i grafi che condividono lo stesso pattern hanno anche lo stesso score. Anche in questo caso, però, affrontare tutte le classi di equivalenza è spesso impossibile. Perciò, in pratica, spesso per ridurre lo spazio di ricerca si ricorre all'utilizzo di euristiche. In certi casi, inoltre, si limita anche lo spazio dei possibili DAG, ad esempio limitandosi al caso in cui essi siano alberi. Si tratta, com'è facilmente intuibile, di approcci che non possono in generale garantire l'ottimalità della soluzione trovata, ma che consentono di trovare una soluzione approssimata in un tempo ragionevole.

Algoritmi score based per l'apprendimento di struttura nel caso di dati completi sono, solo per citare alcuni esempi, GES (Greedy Equivalent Search) (Meek 1997) (Chickering 2002), K2 (Cooper e Herskovits 1992), SC (Sparse Candidate) (Friedman, Nachman e Pe'er 1999), OR (Optimal Reinsertion) (Moore e Wong 2003).

### Approccio constraint based

In questo caso non siamo più interessati a calcolare la probabilità dei dati in base ai vari pattern DAG, ma partiamo dall'ipotesi di conoscere l'insieme delle indipendenze condizionate di una distribuzione di probabilità e trovare un DAG tale per cui la condizione di Markov induca tutte e sole tali indipendenze condizionate. Per prima cosa, si stima l'insieme  $IND_P$  delle indipendenze condizionate della distribuzione  $P$  a partire dai dati. Successivamente, cercheremo un pattern DAG (non potendo distinguere le indipendenze condizionate nei DAG della stessa classe di equivalenza) tale per cui le sue d-separazioni siano le stesse in  $IND_P$ . Questo, però, richiede l'assunzione di faithfulness (o almeno di una sua variante leggermente meno stringente).

I due principali problemi di questo approccio sono la scarsa robustezza, in quanto piccole variazioni nei dati possono cambiare gli esiti dei test di equivalenza ed avere perciò ripercussioni notevoli sulla struttura finale, ed il fatto che tipicamente questi metodi non si rivolgono al caso di dati incompleti.

In questa classe rientrano algoritmi quali PC (dal nome dei due autori Peter Spirtes e Clark Glymour) (Spirtes, Glymour e Scheines 2000), IC (Inductive Causation) (Pearl e Verma 1991), TPDA (Three Phase Dependency Analysis) (Cheng et al. 2002a) e GS (Grow-Shrink) (Margaritis 2003).

Esistono anche metodi, come MMHC (Max-Min Hill Climbing) (Tsamardinos, Brown e Aliferis 2006), che utilizzano approcci ibridi al processo di apprendimento. Quest'ultimo metodo è stato utilizzato nella nostra analisi e verrà descritto in dettaglio in seguito.

### Problemi con la missingness

Per quanto siano presenti in letteratura esempi di apprendimento strutturale con dati incompleti che utilizzano un approccio constraint based, ad es. (Tian, Zang e Lu 2003), la maggior parte dei metodi è invece score based (Friedman 1997), (Friedman 1998), (Borchani, Amor e Khalfallah 2007), (Riggelsen e Feelders 2005), (Riggelsen 2006), (Wong e Guo 2008), (Elidan 2011). Il problema è che, come si è visto, nel caso completo gli score sono decomponibili, mentre questa proprietà viene persa nel passaggio a dati mancanti. Questo implica che la distribuzione a posteriori dei parametri di un modello non è più il prodotto di posteriori indipendenti, impedendoci di rappresentarla in forma chiusa e di fare predizioni esatte usando  $P(X_i | \mathcal{G}, \mathcal{D})$ . Ci vediamo quindi costretti ad approssimare, tipicamente con una stima MAP:

$$P(X_i | \mathcal{G}, \mathcal{D}) \approx P(X_i | \mathcal{G}, \hat{\mathbf{p}})$$

in cui  $\hat{\mathbf{p}}$  è la stima MAP dei parametri del modello, ottenuta ad esempio tramite EM.

A questo punto, dobbiamo calcolare direttamente la posterior complessiva

$$P(\mathcal{D}, \mathcal{G}) = \int P(\mathcal{D} | \mathcal{G}, \mathbf{p}) P(\mathbf{p} | \mathcal{G}) d\mathbf{p}$$

Essendo computazionalmente pesante il calcolo di questo integrale (dato che anch'esso tipicamente non ha forma chiusa), si procede ad approssimazioni, come l'*approssimazione di Laplace*, nel caso di grandi campioni. Essa stima l'integrale eseguendo un fit tramite gaussiane in un intorno del valore MAP dei parametri, il che pone questo problema: per ogni modello, occorre per prima cosa trovare i valori MAP dei suoi parametri, altrimenti non è possibile calcolarne lo score. Chiaramente, se lo spazio dei modelli è ampio, questo approccio richiede un costo computazionale molto alto, in quanto ogni modifica del DAG richiede una nuova stima dei parametri. Vedremo nella prossima sezione alcune delle soluzioni che sono state proposte in letteratura per ovviare a questo inconveniente.

Esiste un'ulteriore casistica, non presa in esame da noi ma che può comunque essere risolta da alcuni algoritmi oggetto di studio, cioè la presenza di variabili nascoste (*hidden*). In questo caso, ad essere incognite non sono soltanto tutti i dati relativi a tali variabili, ma la loro stessa presenza. Questo introduce un altro livello di difficoltà al problema: la presenza di una variabile nascosta può semplificare in maniera notevole il modello ed aiutare molto a spiegare i dati; d'altra parte, sarebbe necessario individuare *quante* variabili nascoste possono essere utili, senza però introdurne in eccesso (ed eventualmente considerandone la plausibilità nel mondo reale). Alcuni algoritmi per il trattamento di dati incompleti possono studiare anche il caso di variabili nascoste, ma tipicamente richiedono che l'utente indichi quante sono.

## 2.2 Stato dell'arte

Di seguito è riportato l'elenco dei metodi che saranno presi in esame. Essi non comprendono tutti gli algoritmi esistenti per trattare l'apprendimento di struttura delle reti bayesiane a partire da dati incompleti, ma sono gli unici per cui sia stato possibile reperire un'implementazione funzionante di pubblico accesso.

- Metodologie generiche (Complete Cases Analysis, Available Cases Analysis, k-nn)
- MMHC (Max Min Hill Climbing)
- MWST (Maximum Weight Spanning Tree)
- SEM (Structural Expectation Maximization)
- GES-EM (Greedy Equivalence Search - Expectation Maximization)

Si riporta qui la notazione utilizzata nel proseguio, che comunque è la stessa già presentata più volte in questa tesi:

- $\mathcal{D} = \{X_1, \dots, X_M\}$  indica un dataset di taglia  $M$ , di cui  $X_i$  sono le variabili;
- $X_i^{(j)}$  è l'istanza della variabile  $i$ -esima nel caso  $j$ -esimo del dataset;
- $\mathcal{D}_o$  rappresenta l'insieme della variabili osservate,  $\mathcal{D}_m$  l'insieme di quelle mancanti. Chiaramente,  $\mathcal{D} = \{\mathcal{D}_o, \mathcal{D}_m\}$ ;
- $\mathcal{G}^{(t)}$  è il modello all'iterazione  $t$ -esima. Quando sia chiaro dal contesto, per semplicità, lo indicheremo usualmente come  $\mathcal{G}$ ;
- $\mathbf{G}$  indica l'insieme dei modelli in esame, ovvero  $\mathcal{G}^{(t)} \in \mathbf{G} \quad \forall t$ .

### 2.2.1 Metodologie generiche

Esistono molte tecniche per trattare dati incompleti che non sono state sviluppate esplicitamente per le reti bayesiane e si prestano all'uso in qualsiasi ambito ci si trovi ad avere a che fare con tale problema. Una panoramica sui metodi per lavorare su dati mancanti -pur se non riferita al caso delle BN- si può trovare nella chiara review (Garcia-Laencina e Sancho-Gómez 2010).

Abbiamo scelto di accoppiare tre di queste metodologie generiche all'algoritmo MMHC, che impara la struttura di reti bayesiane a partire da dati completi (Tsamardinos, Brown e Aliferis 2006) e che descriveremo subito dopo.

**Complete Cases Analysis** Consiste semplicemente nell'eliminare dall'analisi i record che presentino anche un solo elemento incompleto. I vantaggi sono evidenti: è immediato da implementare e consente di utilizzare i metodi per dati completi. Gli svantaggi riguardano la massima percentuale di incompletezza per cui è sensato utilizzarla (che in genere si ritiene dovrebbe essere minore del 5%) ed il fatto che non è poi possibile neanche eseguire test su dati incompleti (perchè non vengono presi in considerazione). Nel seguito la indicheremo anche con l'acronimo *CCA*.

**Available Cases Analysis** In questo caso si utilizzano tutti i dati disponibili nel dataset, anche se un record ha delle voci mancanti. Più precisamente, per un generico nodo  $X$  si prende in considerazione tutta la sua famiglia (cioè tutti i nodi genitori e figli di  $X$ ) e si accetta ogni record che sia completo in tale insieme, anche qualora il record in questione abbia altri nodi i cui valori siano assenti. Si tratta anche in questo caso di un metodo semplice e, oltre tutto, esso generalmente si comporta meglio rispetto alla

CCA in quanto utilizza tutti i dati a disposizione. D'altra parte, spesso non è possibile applicare direttamente gli algoritmi per dati completi e, soprattutto, il numero di casi analizzati non è lo stesso per i vari elementi dei vettori di dati. Ad esso faremo riferimento anche con l'acronimo *ACA*.

**k-nn** (contrazione per “k-Nearest Neighbor”) Il k-nn rientra nel ventaglio dei cosiddetti metodi di imputazione (o *imputation*), che consistono nello stimare il valore del dato mancante secondo qualche criterio e “riempire il buco” (imputare, appunto) con tale stima, in modo da ottenere un dataset completo. Il criterio adottato dal k-nn consiste nel cercare  $k$  vettori tra i dati completi (o che almeno siano completi per il dato mancante nel record attuale) che minimizzino una funzione distanza rispetto al record del dato mancante. Nella nostra implementazione si è utilizzata la *Heterogeneous Euclidean Overlap Metric (HEOM)*:

$$D(\mathbf{x}_a, \mathbf{x}_b) = \sqrt{\sum_{i=1}^n D_i(x_{ia}, x_{ib})^2}$$

dove  $\mathbf{x}_a$  e  $\mathbf{x}_b$  sono due vettori di dati,  $x_{ia}$  e  $x_{ib}$  ne rappresentano le  $i$ -esime componenti e  $D_i(x_{ia}, x_{ib})$  la distanza tra di esse, definita da:

$$D_i(x_{ia}, x_{ib}) = \begin{cases} 1, & (1 - m_{ia})(1 - m_{ib}) = 0, \\ D_0(x_{ia}, x_{ib}), & \text{se } x_i \text{ è un valore discreto,} \\ D_N(x_{ia}, x_{ib}), & \text{se } x_i \text{ è un valore numerico.} \end{cases}$$

in cui, come sopra definito,  $m_{ij}$  indica che nel vettore  $x_j$  l'elemento  $i$ -esimo è mancante. In pratica,  $D_i(x_{ia}, x_{ib})$  varia tra 0 ed 1, in modo che

- se almeno uno dei due vettori ha il dato  $i$ -esimo mancante, la distanza tra gli elementi  $i$ -esimi è massima, cioè vale 1;
- se l'attributo è discreto, intendendo con ciò che esso rappresenta una categoria e non un numero, la distanza  $D_0$  vale 0 se i due attributi sono uguali, 1 altrimenti;
- se l'attributo è numerico, la distanza  $D_N$  è definita come:

$$D_N(x_{ia}, x_{ib}) = \frac{|x_{ia} - x_{ib}|}{\max(x_i) - \min(x_i)}$$

dove  $\max(x_i)$  e  $\min(x_i)$  indicano rispettivamente il massimo ed il minimo valore osservato nel training set per l'attributo numerico  $x_i$ .

A questo punto, occorre definire in che modo si eseguirà l'imputazione, basandosi sui  $k$  dati più simili trovati. Nel nostro caso, abbiamo scelto il valore mediano per gli attributi numerici, la moda per quelli categorici.

I vantaggi del k-nn consistono spesso in buoni risultati, generalmente migliori rispetto ad altri metodi di imputazione; i principali svantaggi sono che i vicini vanno cercati nell'intero dataset, con possibile notevole costo computazionale, e il fatto che si introduce un parametro ( $k$ ) di cui è necessario fare il tuning.

### 2.2.2 MMHC (Max Min Hill Climbing)

L'algoritmo Max-Min Hill-Climbing (MMHC) è un algoritmo per l'apprendimento strutturale delle BN nel caso di dati completi, presentato in (Tsamardinos, Brown e Aliferis 2006). Esso si sviluppa in due fasi: per prima cosa impara lo scheletro della rete e poi, al passo successivo, orienta gli archi per mezzo di una ricerca con hill-climbing greedy guidata da uno score bayesiano. Nel loro articolo, gli Autori eseguono una vasta comparazione rispetto agli altri metodi per l'apprendimento con dati completi e mostrano come MMHC sia contemporaneamente molto veloce, scalabile per grandi reti e garantisca buone performance rispetto a SHD e BDeu. Questo è un metodo che mostra un'ottima qualità di ricostruzione delle reti ed ha il vantaggio di essere veloce e scalabile, anche se esistono altri metodi che in certe condizioni hanno performance superiori (si veda ad es. la comparazione in (Badaloni, Sambo e Venco In press)). Lo abbiamo utilizzato sui dati completi dei nostri dataset per valutare i limiti superiori di qualità a cui è ragionevole aspettarsi che un metodo per dati incompleti possa tendere. Inoltre, l'abbiamo abbinato alle metodiche generali appena descritte (le quali, ognuna a suo modo, si riconducono a dataset completi) per imparare la struttura di una BN anche in caso di dati con missingness.

#### MMPC (Max-Min Parents and Children)

MMPC è un algoritmo di ricerca locale, già presentato in (Tsamardinos, Aliferis e Statnikov 2003), utilizzato per ricostruire lo scheletro della rete. Esso restituisce per ogni nodo  $T$  l'insieme  $PC_T^{\mathcal{G}}$  dei suoi nodi genitori e figli, insieme che non cambia per tutte le reti con diverso DAG ma che siano faithful alla stessa distribuzione  $P$ . In altre parole, un nodo  $X \in PC_T^{\mathcal{G}}$  può essere genitore di  $T$  per un certo  $\mathcal{G}$  e figlio di  $T$  in un diverso  $\mathcal{G}'$ , ma ciò non altera  $PC_T^{\mathcal{G}}$ .

Per prima cosa, viene presentata una versione semplificata di MMPC, chiamata MMPC0, che, assumendo la faithfulness, non restituisce falsi negativi ma può potenzialmente restituire falsi positivi. Questa funzione invoca una funzione per testare l'indipendenza  $Ind(X, T | \mathbb{Z})$  tra le variabili  $X$  e  $T$  dato l'insieme di nodi  $\mathbb{Z}$ , sulla base del dataset di training  $\mathcal{D}$ . La funzione  $Assoc(X, T | \mathbb{Z})$  è una stima della forza della dipendenza tra  $X$  e  $T$  dato  $\mathbb{Z}$ . Per i dettagli implementativi di  $Assoc$  si rimanda all'articolo originale. AAAAAAAAAAAAAA  $MinAssoc(X, T | \mathbb{Z}) \equiv \min_{s \in \mathbb{Z}} Assoc(X, T | s)$ . MMPC utilizza una strategia in due fasi per ottenere  $PC_T^{\mathcal{G}}$ :

**Forward phase** Le variabili sono aggiunte per ottenere un candidato  $PC_T^{\mathcal{G}}$ , chiamato  $CPC$ , con l'uso dell'euristica Max-Min, che prevede di selezionare la variabile che massimizza l'associazione minima con  $T$  relativamente a  $CPC$ .

**Backward phase** Vengono rimossi i falsi positivi introdotti nella prima fase, testando  $Ind(X, T | \mathbb{S})$  per alcuni sottoinsiemi  $\mathbb{S} \subseteq CPC$ . Se i nodi sono condizionatamente indipendenti,  $X$  viene rimosso dalla  $CPC$ .

Lo pseudocodice di MMPC0 è riportato in tabella Algorithm 1.

Tsamardinos et al. mostrano come, in alcuni casi, MMPC0 possa restituire falsi positivi. Gli Autori spiegano che, in caso di falsi positivi, si ha una rottura

---

**Algorithm 1** Algoritmo MMPC0

---

```

1: procedure MMPC0( $T, \mathcal{D}$ )
2:   Input: nodo  $T$ , dataset  $\mathcal{D}$ .
3:   Output: genitori e figli di  $T$  in ogni BN per cui valga la faithul rispetto
   alla distribuzione dei dati.
4:
5:   % Fase 1: Forward
6:    $CPC = \emptyset$ 
7:   repeat
8:      $\langle F, assocF \rangle = MaxMinHeuristic(T, CPC)$ 
9:     if  $assocF \neq 0$  then
10:       $CPC = CPC \cup F$ 
11:     end if
12:   until  $CPC$  non è cambiata
13:
14:   % Fase 2: Backward
15:   for  $\forall X \in CPC$  do
16:     if  $\exists \mathbb{S} \subseteq CPC$  t.c.  $Ind(X, T | \mathbb{S})$  then
17:        $CPC = CPC \setminus \{X\}$ 
18:     end if
19:   end for
20:
21:   restituisci  $CPC$ 
22: end procedure
23:
24: procedure MAXMINHEURISTIC( $T, CPC$ )
25:   Input: nodo  $T$ ,  $CPC$ .
26:   Output: il valore massimo dell'associazione minima fra tutte le variabili
   con  $T$  relativamente a  $CPC$  e la variabile che assume tale massimo.
27:
28:    $assocF = \max_{X \in \mathcal{V}} MinAssoc(X, T | CPC)$ 
29:    $F = \arg \max_{X \in \mathcal{V}} MinAssoc(X, T | CPC)$ 
30:   restituisci  $\langle F, assocF \rangle$ 
31: end procedure

```

---

della simmetria negli insiemi  $PC_T^G$  (evidentemente: di norma, se  $X$  è genitore o figlio di  $T$ , anche  $T$  sarà genitore o figlio per  $X$ ). Il problema viene risolto riportando l'algoritmo MMPC, che richiama MMPC0, ma poi controlla che, se  $X$  appartiene a  $PC_T^G$ , allora anche  $T$  appartenga a  $PC_X^G$ . Qualora ciò non avvenga, significa che  $X$  in realtà non appartiene a  $PC_T^G$  e va pertanto eliminata dall'insieme.

---

**Algorithm 2** Algoritmo MMPC
 

---

```

1: procedure MMPC( $T, \mathcal{D}$ )
2:    $CPC = MMPC0(T, \mathcal{D})$ 
3:   for  $\forall X \in CPC$  do
4:     if  $T \notin MMPC0(X, \mathcal{D})$  then
5:        $CPC = CPC \setminus X$ 
6:     end if
7:   end for
8:
9:   restituisci  $CPC$ 
10: end procedure

```

---

Infine, in tabella Algorithm 3 si riporta l'algoritmo completo di MMHC. Esso identifica tramite MMPC l'insieme di genitori e figli per ogni nodo della rete ed infine esegue una ricerca hill-climbing greedy nello spazio delle reti bayesiane (B-spazio). Tale ricerca parte da un grafo vuoto e seleziona l'aggiunta, rimozione o inversione di un arco che genera il maggior incremento dello score (BDeu). La differenza rispetto ad una ricerca greedy standard è che l'aggiunta di un arco è permessa solo se tale arco è stato precedentemente individuato da MMPC.

---

**Algorithm 3** Algoritmo MMHC
 

---

```

1: procedure MMHC( $\mathcal{D}$ )
2:   Input: dataset  $\mathcal{D}$ 
3:   Output: un DAG con le variabili in  $\mathcal{D}$ 
4:
5:   % Restrizione
6:   for  $\forall X \in \mathcal{V}$  do
7:      $PC_X = MMPC(X, \mathcal{D})$ 
8:   end for
9:
10:  % Ricerca
11:  Partendo da un grafo vuoto, esegui una ricerca hill-climbing greedy con
    gli operatori add-edge, delete-edge, reverse-edge. Usa add-edge  $Y \rightarrow X$  solo
    se  $Y \in PC_X$ .
12:
13:  Restituisci il DAG con lo score più alto trovato.
14: end procedure

```

---

### 2.2.3 SEM (Structural Expectation Maximization)

L'algoritmo SEM (Structural Expectation Maximization) costituisce, *de facto*, il gold standard per il problema che stiamo esaminando. Per quanto sia stato sviluppato quasi vent'anni fa, (Friedman 1997), (Friedman 1998), infatti, esso viene tipicamente considerato il riferimento con cui confrontarsi quando si presentano nuovi algoritmi sull'apprendimento strutturale di BN in presenza di missingness.

Questo metodo nasce come estensione dell'EM (Dempster, Laird e Rubin 1977), aggiungendo un passaggio per la ricerca della struttura. L'algoritmo per il caso di score tipo BIC o MDL, o comunque nella forma

$$S_O(\mathcal{G}, \mathbf{p}) = \log P(\mathcal{D}_o : \mathcal{G}, \mathbf{p}) - Pen(\mathcal{G}, \mathbf{p}, \mathcal{D}_o) \quad (2.16)$$

cioè come log-likelihood dei dati a cui è sottratto un valore di penalità per scoraggiare modelli troppo complicati, viene qui riportato, come presentato in (Friedman 1997) sotto il nome di AMS-EM. Il punto fondamentale è che, non avendo a disposizione dati completi, non si cerca di massimizzare direttamente lo score ma, data una certa stima  $(\mathcal{G}^*, \mathbf{p}^*)$ , la sua aspettazione:

$$Q(\mathcal{G}, \mathbf{p} : \mathcal{G}^*, \mathbf{p}^*) = E[\log P(\mathcal{D}_o, \mathbf{h} : \mathcal{G}, \mathbf{p}) - Pen(\mathcal{G}, \mathbf{p}, \mathcal{D}_o)]$$

dove il valore atteso è calcolato sui valori  $\mathbf{h} \in \mathcal{D}_m$  secondo la distribuzione  $P(\mathbf{h} | \mathcal{D}_o : \mathcal{G}^*, \mathbf{p}^*)$ .

---

#### Algorithm 4 Algoritmo AMS-EM

---

- 1: **procedure** AMS-EM
  - 2:   Scegliere casualmente  $\mathcal{G}^{(0)}$  e  $\mathbf{p}^{(0,0)}$
  - 3:   Ciclo per  $n = 0, 1, \dots$  fino a convergenza
  - 4:     Ciclo per  $l = 0, 1, \dots$  fino a convergenza o  $l = l_{max}$
  - 5:      $\mathbf{p}^{(n,l+1)} = \arg \max_{\mathbf{p}} Q(\mathcal{G}^{(n)}, \mathbf{p} : \mathcal{G}^{(n)}, \mathbf{p}^{(n,l)})$
  - 6:     Trova un modello  $\mathcal{G}^{(n+1)}$  che massimizzi  $Q(\cdot : \mathcal{G}^{(n)}, \mathbf{p}^{(n,l)})$
  - 7:      $\mathbf{p}^{(n+1,0)} = \arg \max_{\mathbf{p}} Q(\mathcal{G}^{(n+1)}, \mathbf{p} : \mathcal{G}^{(n)}, \mathbf{p}^{(n,l)})$
  - 8: **end procedure**
- 

AMS-EM sta per *Alternating Model Selection-EM*, nome che descrive perfettamente la filosofia dell'algoritmo: infatti, all'interno del primo ciclo, alterna costantemente il tentativo di trovare i migliori parametri possibili relativi alla struttura attuale (procedimento computazionalmente più leggero), fino a convergenza o al raggiungimento di un numero di passi EM predefinito, e soltanto successivamente tenta di cambiare il modello (passo che ha un peso computazionale maggiore).

Friedman ha presentato l'anno successivo una versione di SEM per score di tipo bayesiano (Friedman 1998). In questo caso assumiamo di avere un prior relativo ai modelli e all'assegnazione dei parametri per ogni modello. Noi vorremmo calcolare una stima MAP

$$P(\mathcal{D} | \mathcal{G}_h)$$

ma, in caso di dati mancanti, non possiamo farlo efficientemente. Supponiamo allora di poter stimare la verosimiglianza dei dati completi (ad esempio nel caso di modelli decomponibili)

$$P(\mathcal{D}_m, \mathcal{D}_o | \mathcal{G}_h)$$

**Algorithm 5** Algoritmo SEM-Bayesiano

---

```

1: procedure SEM-BAYESIANO( $\mathcal{G}^{(0)}, \mathbf{o}$ )
2:   Ciclo per  $n = 0, 1, \dots$  fino a convergenza
3:
4:   Calcola la posterior  $P(\mathbf{p}^{\mathcal{G}^{(n)}} | \mathcal{G}_h^{(n)}, \mathbf{o})$ 
5:
6:   E-step: Per ogni  $\mathcal{G}$ , calcola
7:

```

$$\begin{aligned}
Q(\mathcal{G} : \mathcal{G}^{(n)}) &= E[\log P(\mathcal{D}_m, \mathbf{o}) | \mathcal{G}_h^{(n)}, \mathbf{o}] \\
&= \sum_{\mathbf{h} \in \mathcal{D}_m} P(\mathbf{h} | \mathbf{o}, \mathcal{G}_h^{(n)}) \log P(\mathbf{h}, \mathbf{o}, \mathcal{G}_h)
\end{aligned}$$

```

8:   M-step: Scegli  $\mathcal{G}^{(n+1)}$  che massimizzi  $Q(\mathcal{G} : \mathcal{G}^{(n)})$ 
9:
10:  if  $Q(\mathcal{G}^{(n)} : \mathcal{G}^{(n)}) = Q(\mathcal{G}^{(n+1)} : \mathcal{G}^{(n)})$  then
11:    restituisci  $\mathcal{G}^{(n)}$ 
12:  end if
13: end procedure

```

---

In entrambi i casi, non è necessario massimizzare il valore atteso dello score su tutti i modelli, in quanto Friedman ha dimostrato che per avvicinarsi alla convergenza è sufficiente scegliere  $\mathcal{G}^{(n+1)}$  tale da aumentare il valore dello score:

$$Q(\mathcal{G}^{(n+1)} : \mathcal{G}^{(n)}) > Q(\mathcal{G}^{(n)} : \mathcal{G}^{(n)})$$

Inoltre, poichè entrambe le versioni tendono a fermarsi in massimi locali, si eseguirà più volte l'algoritmo partendo da strutture e parametri iniziali diversi (ad es., scelte casualmente), scegliendo infine il risultato con lo score più alto.

### 2.2.4 MWST-EM (Maximum Weight Spanning Tree - Expectation Maximization)

Una fusione dell'algoritmo proposto da (Friedman 1997), ossia AMS-EM con BIC score, e di quello proposto in (Chow e Liu 1968) è stata proposta da Leray e Francois (Leray e Francois 2005). Gli Autori hanno cioè ristretto l'insieme  $\mathbf{G}$  in cui AMS-EM ricerca il DAG migliore all'insieme  $\mathbf{T}$  degli alberi, per cercare il migliore Maximum Weight Spanning Tree. Un MWST è un albero che connette tutti i nodi di  $\mathcal{V}$  massimizzando una certa funzione obiettivo (o, equivalentemente, minimizzando una determinata funzione costo). Com'è ovvio, si tratta di una semplificazione dello spazio di ricerca che tendenzialmente ci aspettiamo non individui la struttura reale, dato che questa spesso non è un albero. Tuttavia, essa consente di individuare una soluzione in tempo polinomiale nel numero di nodi del grafo. Inoltre, nel loro articolo gli Autori mostrano dei risultati sperimentali in cui, inizializzando AMS-EM con MWST-EM, la struttura trovata da SEM (quindi nello spazio  $\mathbf{G}$  dei DAG) è spesso migliore rispetto a quella ottenuta con l'inizializzazione standard (rete casuale) di SEM. Si riporta l'algoritmo nella tabella Algorithm 6.

**Algorithm 6** Algoritmo MWST-EM

---

```

1: procedure MWST-EM
2:   Inizializzazione: finito=falso,  $i = 0$ 
3:   Scelta casuale o euristica dell'albero e dei parametri iniziali  $(\mathcal{T}^{(0)}, \Theta^{(0,0)})$ 
4:   repeat
5:      $j=0$ 
6:     repeat
7:        $\Theta^{(i,j+1)} = \arg \max_{\Theta} Q(\mathcal{T}^{(i)}, \Theta : \mathcal{T}^{(i)}, \Theta^{(i,j)})$ 
8:        $j = j + 1$ 
9:     until convergenza  $(\Theta^{(i,j)} \rightarrow \Theta^{(i,j^o)})$ 
10:    if  $i = 0$  o  $| Q(\mathcal{T}^{(i)}, \Theta^{(i,j^o)} : \mathcal{T}^{(i-1)}, \Theta^{(i-1,j^o)}) - Q(\mathcal{T}^{(i-1)}, \Theta^{(i-1,j^o)} : \mathcal{T}^{(i-1)}, \Theta^{(i-1,j^o)}) > \epsilon$  then
11:       $\mathcal{T}^{(i+1)} = \arg \max_{\mathcal{T} \in \mathbf{T}} Q(\mathcal{T}, \cdot : \mathcal{T}^{(i)}, \Theta^{(i,j^o)})$ 
12:       $i = i + 1$ 
13:
14:    else
15:      finito=vero
16:    end if
17:  until finito=true
18: end procedure

```

---

L'unica differenza di questo algoritmo rispetto ad AMS-EM è nell'equazione che decide la migliore struttura per l'iterazione successiva (riga 11), in quanto le parti precedenti riguardano un semplice EM parametrico relativo alla struttura dell'iterazione  $i$ -esima. La funzione di similarità (cioè lo score) tra due nodi usata dall'algoritmo MWST originario era basata su variante dello score BIC nel caso il nodo  $X_i$  sia collegato ad  $X_j$  o no, e può essere espressa come matrice simmetrica:

$$[M_{ij}] = [bic(X_i, pa_i = X_j, \Theta_{X_i|X_j}) - bic(X_i, pa_i = \emptyset, \Theta_{X_i})]$$

in cui  $bic$  indica il BIC locale già definito. A questo punto, un algoritmo per MWST come quello di Kruskal (Kruskal 1956) applicato alla matrice  $M$  restituiva il miglior albero  $\mathcal{T}$  che ottimizzava il BIC globale.

Nell'algoritmo in esame, l'adattamento al caso di dati incompleti avviene sostituendo i BIC locali con la loro aspettazione relativamente ad una densità di probabilità dei dati mancanti  $P(\mathcal{D}_m | \mathcal{T}^*, \Theta^*)$ , ossia:

$$[M_{ij}^Q] = [Q^{bic}(X_i, pa_i = X_j, \Theta_{X_i|X_j} : \mathcal{T}^*, \Theta^*) - Q^{bic}(X_i, pa_i = \emptyset, \Theta_{X_i} : \mathcal{T}^*, \Theta^*)]$$

Allo stesso modo, un algoritmo per trovare un MWST può essere eseguito sulla matrice  $M^Q$ , permettendoci di ottenere il miglior albero  $\mathcal{T}$  che massimizza la funzione  $Q^{BIC}$ , già vista per AMS-EM.

### 2.2.5 GES-EM (Greedy Equivalence Search - Expectation Maximization)

Il Greedy Equivalence Search-Expectation Maximization è un algoritmo per dati incompleti proposto in (Borchani, Amor e Khalfallah 2007). Si tratta della variante per dati mancanti dell'algoritmo GES presentato da (Chickering 2002),

in cui lo spazio di ricerca non è quello dei DAG (chiamato anche *B-spazio*), ma quello delle classi di equivalenza dei DAG, detto *E-spazio*. Il vantaggio di questa scelta è che si riduce il numero di strutture disponibili, che indica una scelta concettualmente simile a quella di cercare soltanto tra gli alberi, senza però perdere in qualità della struttura. Infatti, muoversi nell'E-spazio significa comunque cercare tra tutti i DAG, con la sola differenza, rispetto ad una ricerca greedy nel B-spazio, che le strutture equivalenti (che hanno tutte lo stesso score) vengono prese in esame soltanto una volta.

Il GES-EM si compone di due passi iterativi che verranno ora descritti. Lo score utilizzato, a differenza del metodo GES a cui esso si ispira, è il BIC. Inoltre, la principale differenza rispetto a GES è l'introduzione di un'alternanza tra iterazioni di tipo EM per ottimizzare i parametri della classe e iterazioni per la ricerca della struttura, che hanno l'obiettivo di individuare i possibili vicini della classe di equivalenza (applicando GES).

**Inizializzazione** L'algoritmo inizia con una classe di equivalenza  $\xi^{(0)}$  il cui unico elemento è l'insieme vuoto  $\mathcal{G}^{(0)}$ . I parametri sono inizializzati casualmente a  $\Theta^{(0)}$ .

**Fase 1 (inserzione di un singolo arco)** Nella prima fase, ad ogni iterazione  $i$ , con  $i = 0, 1, \dots, i_{MAX}$ , si esegue l'algoritmo EM fino a convergenza, partendo dallo stato attuale  $(\mathcal{G}^{(i)}, \Theta^{(i)})$  e trovando i parametri ottimizzati  $\Theta'^{(i)}$ . Successivamente, applica tutti le operazioni di inserimento *Insert* ad  $\xi^{(i)}$  per ottenere l'insieme  $P^+(\xi^{(i)})$  dei suoi vicini. Ogni vicino è rappresentato da un PDAG  $\mathcal{P} \in P^+(\xi^{(i)})$  che non è necessariamente completo e perciò va convertito nel corrispondente DAG  $\mathcal{G}$  secondo la procedura PDAG-to-DAG presentata in (Dor e Tarsi 1992), la quale verifica la correttezza dell'operazione di inserimento appena eseguita, accettando o rifiutando il corrispondente PDAG  $\mathcal{P}$  (e, di conseguenza,  $\mathcal{G}$ ).

Si calcola quindi il BIC per ogni estensione consistente  $e$ , se questo è maggiore rispetto allo score dell'iterazione precedente, si genera la nuova classe di equivalenza. Se ciò non accade, significa che si è raggiunto (e superato) un massimo locale e si passa alla fase successiva a partire dal risultato dell'iterazione precedente.

In tabella Algorithm 7 è riportato l'algoritmo per questa parte.

**Fase 2 (delezione di un singolo arco)** La seconda fase è del tutto simile alla prima, ad eccezione del fatto che si sostituiscono tutte le operazioni di *Insert* con operazioni di *Delete*, ossia di cancellazione di singoli archi.

Nelle valutazioni sperimentali, Borchani et al. ottengono un metodo consistentemente migliore e più veloce rispetto ad AMS-EM. Si mette però in evidenza che GES, il metodo per dati completi a cui le Autrici si ispirano e che viene utilizzato dentro GES-EM, si dimostra, sia nell'articolo originale di Chickering (Chickering 2002), sia nella comparazione di Tsamardinos et al. (Tsamardinos, Brown e Aliferis 2006), estremamente più lento rispetto ai metodi concorrenti, pur se con risultati di buona qualità.

**Algorithm 7** Algoritmo GES-EM - Inserzione

---

```

1: procedure INSERZIONE
2:    $i \leftarrow 0$ ;  $Convergenza \leftarrow falso$ ;  $Stop \leftarrow falso$ ;
3:   while (not  $Convergenza$ ) e (not  $Stop$ ) do
4:      $\Theta^{(i)} \leftarrow EM(\mathcal{G}^{(i)}, \Theta^{(i)})$ ;
5:     Genera  $P^+(\xi^{(i)})$  applicando tutti gli operatori Insert validi per  $\xi^{(i)}$ ;
6:      $G^+(\xi^{(i)}) \leftarrow PDAG\text{-to-DAG}(P^+(\xi^{(i)}))$ ;
7:     if  $G^+(\xi^{(i)}) \neq \emptyset$  then
8:        $(\mathcal{G}^{(i+1)}, \Theta^{(i+1)}) \leftarrow \arg \max Q(\mathcal{G}, : \mathcal{G}^{(i)}, \Theta^{(i)})$ ;
9:       if  $Q(\mathcal{G}^{(i+1)}, \Theta^{(i+1)} : \mathcal{G}^{(i)}, \Theta^{(i)}) > Q(\mathcal{G}^{(i)}, \Theta^{(i)} : \mathcal{G}^{(i)}, \Theta^{(i)})$  then
10:         $\xi^{(i+1)} \leftarrow DAG\text{-to-CPDAG}(\mathcal{G}^{(i+1)})$ ;
11:         $i \leftarrow i + 1$ ;
12:       else
13:         $Convergenza \leftarrow vero$ ;
14:       end if
15:     else
16:        $Stop \leftarrow vero$ ;
17:     end if
18:   end while
19: end procedure

```

---

**2.2.6 Ulteriori algoritmi**

In questa sezione si illustreranno velocemente gli altri algoritmi per l'apprendimento strutturale di reti bayesiane a partire da dati incompleti che erano stati presi in considerazione, ma per cui non sono disponibili implementazioni di pubblico accesso.

**emc4** Questo metodo, proposto in (Riggelsen e Feelders 2005), ha come obiettivo di ottenere la distribuzione a posteriori dei modelli, in base ai dati osservati. Ciò viene fatto approssimando una distribuzione predittiva per i dati mancanti.

**Markov Blanket Predictor** Algoritmo presentato in (Riggelsen 2006) che si propone di valutare i dati mancanti per mezzo di una distribuzione predittiva basata sulla Markov blanket di ogni nodo. A questo punto, si può ricorrere ad un algoritmo di struttura come l'hill-climbing (Kocka e Castelo 2001).

**Estimated Mutual Information** Questo algoritmo constraint-based si propone di fornire una stima della mutua informazione condizionata tra set di variabili. A tal punto, si ricorre all'apprendimento di struttura presentato da (Cheng, Bell e Liu 1997), (Cheng et al. 2002b). Proposto in (Tian, Zang e Lu 2003).

**Evolutionary BN** Si tratta di un metodo evolutivo ibrido utilizzato in combinazione con il classico EM, proposto in (Wong e Guo 2008). I dati vengono completati facendo ricorso anche a SEM e la ricerca della struttura utilizza l'algoritmo per dati completi HEA (Wong e Leung 2004).

**Bagged Structural Expectation Maximization** Questo metodo, proposto in (Elidan 2011), è una variante di SEM che si propone di sfruttare in maniera intelligente il bagging (Breiman 1996), modificando opportunamente la funzione obiettivo di SEM.



# Capitolo 3

## Comparazione

In questo capitolo si illustrerà l'impianto sperimentale utilizzato per confrontare i metodi esposti al Cap. 2 ed i risultati della comparazione effettuata.

### 3.1 Setup sperimentale

In questa sezione vogliamo illustrare il modo in cui è stata condotta la comparazione degli algoritmi presentati al cap. 2. Sono state selezionate sei diverse reti di pubblico dominio da cui sono stati generati i dati dei training set completi. Ad ognuno di questi dataset è stata imposta missingness secondo tutti e tre i meccanismi citati e, per ogni meccanismo, secondo tre diverse probabilità. A quel punto, i quattro (tredici varianti complessive) algoritmi sono stati eseguiti con l'obiettivo di ricostruire le reti di partenza, che fungevano quindi da gold standard. Infine, sono state valutate le seguenti metriche, che verranno descritte in dettaglio in seguito:

- Structural Hamming Distance rispetto alla struttura di partenza;
- Precision rispetto alla struttura di partenza;
- Recall rispetto alla struttura di partenza;
- Tempo di esecuzione.

Si è partiti da reti note in quanto in questo caso la struttura vera è nota e si evita il rischio (possibile se si utilizza come gold standard una rete ricostruita da un determinato algoritmo di riferimento) che metodi che utilizzano score diversi dal metodo di riferimento siano ingiustamente penalizzati. Per lo stesso motivo, si è scelto di valutare esclusivamente la struttura e non altri punteggi quali BIC/MDL, BDeu o simili.

Al meglio delle nostre conoscenze, questo è il primo studio comparativo così vasto nell'ambito dell'apprendimento strutturale di reti bayesiane in caso di dati mancanti. Il tempo-macchina totale (riferito come relativo ad una singola CPU) è di oltre quattro anni, con un numero totale di reti imparate pari a 29700.

Tutte le prove sono state eseguite su (un singolo core di) Intel Xeon E5450, 12 MB cache, 3.00 GHz, 16 GB RAM equipaggiato con Linux. Il tempo massimo che un algoritmo ha avuto per imparare una rete a partire da un dataset è stato pari a 12 ore.

### 3.1.1 Algoritmi

La lista degli algoritmi testati e delle loro varianti è qui riportata:

- Complete Case Analysis + MMHC (a partire da rete vuota)
- Available Case Analysis + MMHC (a partire da rete vuota)
- k-nn + MMHC (a partire da rete vuota)
- SEM (BDeu) (a partire da rete vuota)
- SEM (BIC) (a partire da rete vuota)
- GES-EM (a partire da rete vuota)
- MWST-EM (a partire da rete vuota)
- MWST-EM → Complete Case Analysis + HC (l'hill-climbing sui soli dati completi viene fatto partire dalla rete ottenuta in uscita all'MWST-EM; lo indicheremo come CCA+T)
- MWST-EM → Available Case Analysis + HC (l'hill-climbing sui soli dati disponibili viene fatto partire dalla rete ottenuta in uscita all'MWST-EM; lo indicheremo come ACA+T)
- MWST-EM → k-nn + HC (l'hill-climbing viene fatto partire dalla rete ottenuta in uscita all'MWST-EM; lo indicheremo come k-nn+T)
- MWST-EM → SEM (BDeu) (il SEM viene fatto partire dalla rete ottenuta in uscita all'MWST-EM)
- MWST-EM → SEM (BIC) (il SEM viene fatto partire dalla rete ottenuta in uscita all'MWST-EM)
- MWST-EM → GES-EM (il GES-EM viene fatto partire dalla rete ottenuta in uscita all'MWST-EM; lo indicheremo come GES-EM+T)

Per quanto si tratti di buona parte degli algoritmi che possono definirsi stato dell'arte, l'insieme non è da ritenersi esaustivo.

Nella scelta della specifica implementazione, abbiamo preferito quelle degli autori, qualora fossero pubblicamente disponibili e funzionassero sulla nostra piattaforma sperimentale. In caso contrario, abbiamo implementato direttamente gli algoritmi, avendo cura che seguissero il più possibile le indicazioni degli autori, anche per ciò che riguarda la scelta dei valori di default dei parametri. In particolare:

- per SEM abbiamo utilizzato l'implementazione di Friedman (autore di SEM) e Elidan in LibB, disponibile come eseguibile per Linux e Windows in (<http://compbio.cs.huji.ac.il/LibB/programs.html>);
- per MWST-EM e GES-EM abbiamo fatto ricorso alle implementazioni degli autori presenti nel pacchetto Matlab© BNT-SLP (scaricabile da <http://www.mathworks.com/matlabcentral/fileexchange/13562-structure-learning-package-for-bayes-net-toolbox>) (Leray e Francois 2004), («The BayesNet Toolbox for Matlab»);

- per MMHC abbiamo implementato l'algoritmo nel pacchetto R (Teau et al. 2012) *bnstruct* (<https://github.com/magodellepercussioni/bnstruct>);

Come parametri abbiamo imposto:

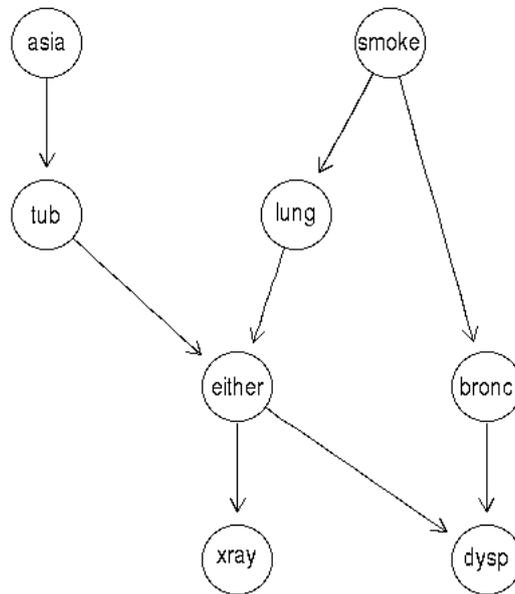
- per k-nn:  $k = 10$ ;
- per MMHC:  $ESS = 1$ , tabu-list di lunghezza 100 e numero massimo di worsening moves pari a 15;
- per SEM:  $ESS = 1$ , modalità hill-climbing (predefinita in LibB), con tabu-list di lunghezza 100 e numero massimo di worsening moves pari a 15;
- per GES-EM: numero massimo di iterazioni per le fasi di inserimento e cancellazione pari a 30 (nel loro articolo, Borchani et al. utilizzano 30 iterazioni sulla rete Asia ed 80 per Insurance, ma abbiamo verificato che per molte reti la convergenza arriva prima che siano raggiunti le 30 iterazioni o il limite temporale massimo che abbiamo imposto). Gli altri parametri sono stati lasciati a default;
- per MWST-EM tutti i parametri sono stati lasciati alle impostazioni base;

Si fa notare che la scelta di testare SEM sia con score di tipo BDeu che BIC è dovuta al fatto che, nel suo articolo del 1998 (Friedman 1998), Friedman mostra come il metodo con BDeu sia più efficiente rispetto all'altro, mentre Leray e Francois (Leray e Francois 2005) eseguono le loro prove utilizzando l'albero ottenuto dal loro algoritmo per inizializzare AMS-EM (cioè SEM con score BIC). E' perciò sembrato corretto testare entrambe le varianti sia a partire da rete vuota che dall'albero MWST.

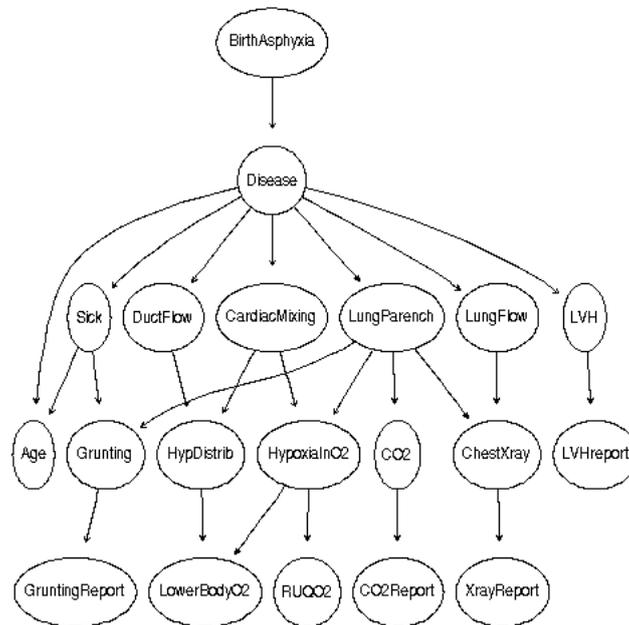
### 3.1.2 Reti

Per la comparazione sono state utilizzate reti pubblicamente disponibili ed in larga parte derivanti da veri sistemi per il supporto alle decisioni utilizzati in ambiti differenti. Esse sono:

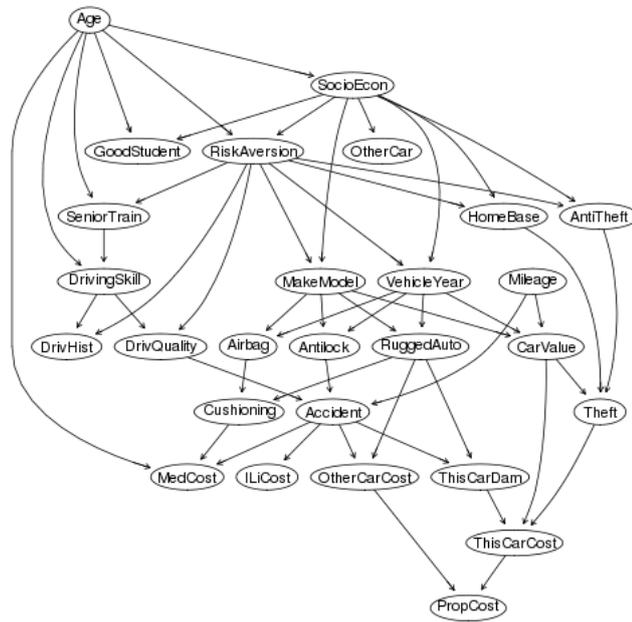
- Asia (Lauritzen e Spiegelhalter 1988)



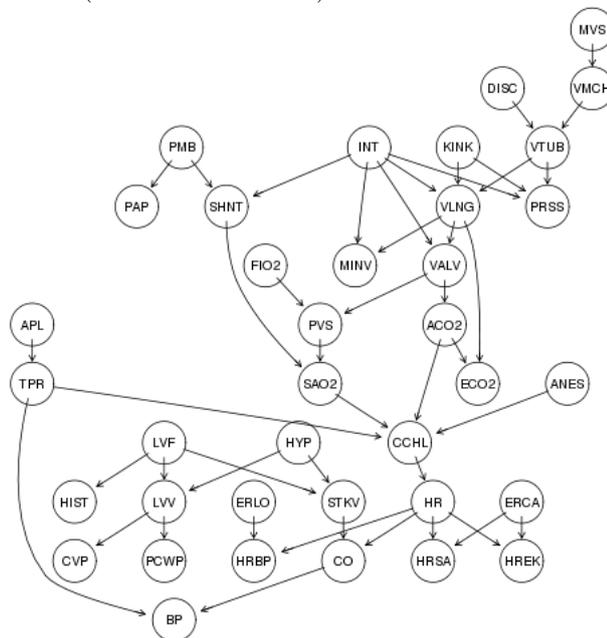
- Child (Cowell et al. 1999)



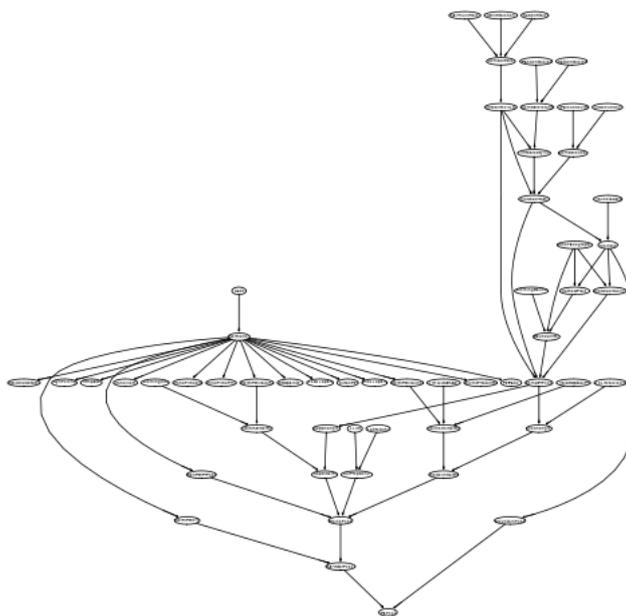
- Insurance (Binder et al. 1997)



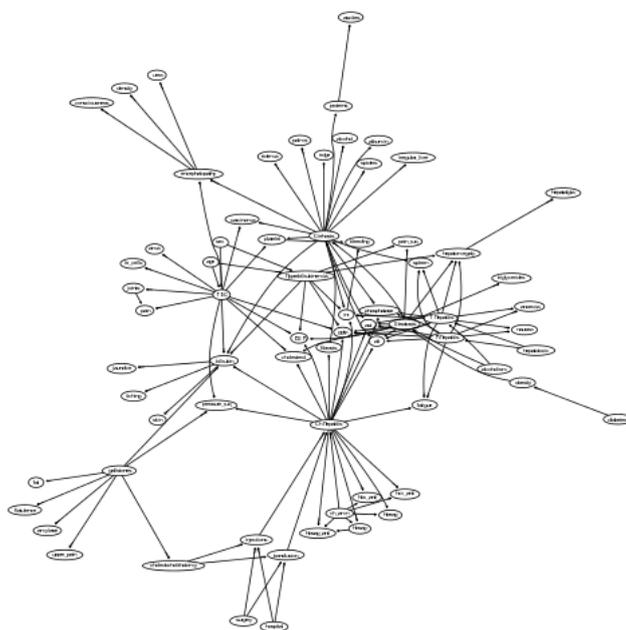
- Alarm (Beinlich et al. 1989)



- Hailfinder (Jensen e Jensen 1996)



- Hepar II (Onisko 2003)



Le immagini sono prese dal repository di pubblico accesso reperibile in [www.bnlearn.com](http://www.bnlearn.com) (Scutari 2009).

In tabella 3.1 sono riportate schematicamente le caratteristiche delle reti oggetto di studio.

Tabella 3.1: Caratteristiche delle reti utilizzate nel nostro studio.

<i>Rete</i>	<i>Num. variabili</i>	<i>Num. archi</i>	<i>In/out-degree massimi</i>	<i>Limiti dominio</i>	<i>Numero parametri</i>
Asia	8	8	2/2	2	18
Child	20	25	2/7	2-6	230
Insurance	27	52	3/7	2-5	984
Alarm	37	46	4/5	2-4	509
Hailfinder	56	66	4/16	2-11	2656
Hepar II	70	123	6/17	2-4	1453

### 3.1.3 Dataset

Per ogni rete siamo partiti da dieci dataset completi di cardinalità 500, dieci di cardinalità 1000 e dieci di cardinalità 5000. Per Child, Insurance, Alarm e Hailfinder, abbiamo utilizzato i dataset disponibili nel sito collegato all'articolo (Tsamardinos, Brown e Aliferis 2006) ([http://www.dsl-lab.org/supplements/mmhc\\_paper/mmhc\\_index.html](http://www.dsl-lab.org/supplements/mmhc_paper/mmhc_index.html)), per Asia e Hepar2 abbiamo generato i dataset tramite Gibbs sampling. Successivamente, abbiamo sottocampionato i dataset di cardinalità 1000 per ottenere altrettanti dataset di taglia 100 e quelli di cardinalità 5000 per ottenerne altri dieci di taglia 200. Come si può vedere, i nostri insiemi di dati includono anche dataset di piccola taglia, in quanto abbiamo ritenuto importante rivolgere la nostra attenzione ai casi in cui si abbia pochi dati, condizione che è, almeno nell'ambito genomico, decisamente più frequente rispetto all'essere in possesso di dataset con taglia oltre il migliaio.

A questo punto, da ognuno dei cinquanta dataset completi ne abbiamo ottenuti nove incompleti, con tre probabilità di missingness diverse per ciascun meccanismo di generazione di dati mancanti. In dettaglio, questi sono i protocolli che abbiamo seguito:

**MCAR** Nel caso di dati *missing completely at random*, un dato deve avere una probabilità di mancare indipendente sia dal valore che esso stesso assume, sia dai valori delle altre variabili.

1. Si fissa l'intervallo  $\eta$  che indica la probabilità desiderata di ottenere dati mancanti ( $\eta \in \{[0, 0.1], [0.1, 0.2], [0.2, 0.3]\}$ ) e ad ogni elemento  $d_{ij}$  del dataset si associa un valore  $\eta_{ij}$  in questo intervallo, secondo una distribuzione uniforme;
2. si assegna un valore  $\theta_{ij}$  nell'intervallo  $\Theta = [0, 1]$  ad ogni elemento del dataset, secondo una distribuzione uniforme;
3. il dato  $d_{ij}$  viene rimosso dal dataset se  $\eta_{ij} \geq \theta_{ij}$ .

**MAR** Se i dati sono mancanti secondo uno schema *missing at random*, la loro assenza dipende dai valori assunti da altre variabili osservabili. Abbiamo imposto questo schema:

1. si divide casualmente l'insieme dei nodi della rete,  $\mathcal{V}$ , in due metà. Una di queste (sia  $\mathcal{V}_c$ ) rappresenterà i nodi condizionanti, l'altra sarà quella dei dati potenzialmente mancanti  $\mathcal{V}_m$ , intendendo con questi termini che i valori delle variabili in  $\mathcal{V}_m$  possono mancare in base

ai valori assunti dai nodi in  $\mathcal{V}_c$ . Se il numero di nodi è dispari, un vertice non rientrerà in alcuna di queste categorie (ed avrà sempre i dati completi);

2. si fissa casualmente una regola per cui ogni nodo in  $\mathcal{V}_c$  induca missingness in uno ed un solo nodo di  $\mathcal{V}_m$ ;
3. per ognuno dei nodi  $V_c \in \mathcal{V}_c$  si suddividono casualmente in due metà i valori che questo può assumere. Chiameremo  $\mathbf{R}_c$  l'insieme dei valori rilevanti per il nodo  $V_c$  e  $\mathbf{N}_c$  l'insieme dei suoi valori non rilevanti, dove questi termini indicano se un certo valore è rilevante o meno per ottenere la missingness nei nodi in  $\mathcal{V}_m$ . Se il numero di valori che il nodo può assumere è dispari, l'insieme  $\mathbf{N}_c$  ha un elemento in più rispetto a  $\mathbf{R}_c$ .

Per chiarire il procedimento seguito finora, se la rete ha quattro vertici  $\mathcal{V} = \{V_1, V_2, V_3, V_4\}$  ognuno dei quali può assumere 6 valori, diciamo  $\mathbf{V} = \{1, 2, 3, 4, 5, 6\}$ , sceglieremo a caso due nodi e li metteremo in  $\mathcal{V}_c$ , poniamo  $\mathcal{V}_c = \{V_2, V_3\}$  (e perciò  $\mathcal{V}_m = \{V_1, V_4\}$ ). Supponiamo poi che la regola ottenuta sia:

$$\begin{array}{ccc} V_2 & \xrightarrow{\text{missingness}} & V_4 \\ V_3 & \xrightarrow{\text{missingness}} & V_1 \end{array}$$

A questo punto, dobbiamo decidere quali valori di  $V_2$  possono causare la mancanza di dati in  $V_4$  (allo stesso modo per  $V_3$  e  $V_1$ ). Ad esempio, fissiamo  $\mathbf{R}_2 = \{1, 5, 6\}$  (e quindi  $\mathbf{N}_2 = \{2, 3, 4\}$ ). Questo significa che, quando  $V_2$  assumerà i valori 1, 5, 6, il dato in  $V_4$  potrebbe mancare, mentre ciò certamente non succederà se  $V_2$  assumerà un valore tra 2, 3 e 4.

4. si fissa l'intervallo  $\eta$  che indica la probabilità desiderata di ottenere dati mancanti nei nodi  $\mathcal{V}_m$  dato un valore rilevante nei corrispondenti nodi in  $\mathcal{V}_c$  ( $\eta \in \{[0, 0.25], [0.25, 0.5], [0.5, 0.75]\}$ ). Ad ogni elemento  $d_{im}$  (elemento del nodo  $V_m \in \mathcal{V}_m$  del record  $i$ -esimo) del dataset si associa un valore  $\eta_{im}$  in questo intervallo, secondo una distribuzione uniforme;
5. si assegna un valore  $\theta_{im}$  nell'intervallo  $\Theta = [0, 1]$  ad ogni elemento dei nodi in  $\mathcal{V}_m$  del dataset, secondo una distribuzione uniforme;
6. il dato  $d_{im}$  di un nodo  $V_m \in \mathcal{V}_m$  viene rimosso dal dataset se il rispettivo (secondo la regola fissata al punto 2) nodo  $V_c \in \mathcal{V}_c$  assume valori in  $\mathbf{R}_c$  e  $\eta_{im} \geq \theta_{im}$ .

**NMAR** Nel caso i dati manchino secondo uno schema *not missing at random*, la loro assenza dipende dal valore che la variabile stessa assume e perciò, come si è già detto, la missingness non è più ignorabile. Strettamente parlando, nessuno degli algoritmi oggetto del nostro confronto assume questo meccanismo (fa eccezione k-nn, in quanto valuta la somiglianza tra record senza ipotesi sul meccanismo di incompletezza). Nondimeno, esso nella realtà è presente e costituisce quindi un caso che deve essere affrontato da un algoritmo che lavori su dati incompleti. Abbiamo perciò ritenuto interessante eseguire test anche in questo senso. Di seguito si riporta il protocollo seguito:

1. si divide casualmente l'insieme dei nodi della rete,  $\mathcal{V}$ , in due metà. Una di queste (sia  $\mathcal{V}_{\text{NMAR}}$ ) rappresenterà i nodi i cui dati sono potenzialmente mancanti, l'altra (che chiameremo  $\mathcal{V}_{\text{compl}}$ ) conterrà tutti i nodi i cui dati saranno sempre completi. Se il numero di nodi è dispari,  $\mathcal{V}_{\text{compl}}$  avrà un elemento in più rispetto a  $\mathcal{V}_{\text{NMAR}}$ ;
2. per ognuno dei nodi  $V_n \in \mathcal{V}_{\text{NMAR}}$  si suddividono casualmente in due metà i valori che questo può assumere. Chiameremo  $\mathbf{R}_n$  l'insieme dei valori rilevanti per il nodo  $V_n$  e  $\mathbf{N}_n$  l'insieme dei suoi valori non rilevanti, dove questi termini indicano se un certo valore è rilevante o meno per ottenere missingness nel nodo stesso. Se il numero di valori che il nodo può assumere è dispari, l'insieme  $\mathbf{N}_n$  ha un elemento in più rispetto a  $\mathbf{R}_n$ .  
Ad esempio, riprendiamo la rete ha quattro vertici  $\mathcal{V} = \{V_1, V_2, V_3, V_4\}$  ognuno con valori  $\mathbf{V} = \{1, 2, 3, 4, 5, 6\}$ . Sceglieremo a caso due nodi e li metteremo in  $\mathcal{V}_{\text{NMAR}}$ , supponiamo  $\mathcal{V}_{\text{NMAR}} = \{V_1, V_3\}$  (e perciò  $\mathcal{V}_{\text{compl}} = \{V_2, V_4\}$ ). A questo punto, per  $V_1$  (e allo stesso modo per  $V_3$ ) inseriremo metà valori in  $\mathbf{R}_1$ , diciamo  $\mathbf{R}_1 = \{2, 3, 6\}$  (e perciò  $\mathbf{N}_1 = \{1, 4, 5\}$ ). Procederemo poi a cancellare dal dataset i dati in  $\mathbf{R}_1$  con una certa probabilità, secondo i punti seguenti;
3. si fissa l'intervallo  $\eta$  che indica la probabilità desiderata di ottenere dati mancanti nei nodi  $V_n \in \mathcal{V}_{\text{NMAR}}$  quando essi assumono un valore rilevante ( $\eta \in \{[0, 0.25], [0.25, 0.5], [0.5, 0.75]\}$ ). Ad ogni elemento  $d_{in}$  (elemento del nodo  $V_n \in \mathcal{V}_{\text{NMAR}}$  del record  $i$ -esimo) del dataset si associa un valore  $\eta_{in}$  in questo intervallo, secondo una distribuzione uniforme;
4. si assegna un valore  $\theta_{in}$  nell'intervallo  $\Theta = [0, 1]$  ad ogni elemento dei nodi in  $\mathcal{V}_{\text{NMAR}}$  del dataset, secondo una distribuzione uniforme;
5. il dato  $d_{in}$  di un nodo  $V_n \in \mathcal{V}_{\text{NMAR}}$  viene rimosso dal dataset se assume valori in  $\mathbf{R}_n$  e  $\eta_{in} \geq \theta_{in}$ .

### 3.1.4 Misure di performance

Per misurare la qualità degli algoritmi abbiamo utilizzato le seguenti metriche:

**Tempo** Per verificare la velocità computazionale, abbiamo scelto di misurare il tempo reale di esecuzione di ogni algoritmo. Chiaramente, si tratta di una scelta che risente sia dell'implementazione sia dell'ambiente di sviluppo in cui questa è stata sviluppata. D'altronde, essa include l'insieme di tutti i calcoli svolti da un'applicazione e fornisce comunque un'utile indicazione di massima della velocità dell'algoritmo.

**Prestazioni** Per valutare la qualità della ricostruzione, ci siamo limitati a metriche relative alla struttura delle reti in quanto altri punteggi quali BIC, BDeu o similari tenderebbero a favorire quegli algoritmi che su di esse si basano, a scapito di quelli che utilizzano score diversi.

- La principale misura che abbiamo utilizzato è la **Structural Hamming Distance (SHD)**, introdotta in (Tsamardinos, Brown e Aliferis 2006), che richiede di convertire i due DAG a confronto in PDAG e dà

un punto di penalità per ognuna delle seguenti operazioni necessarie per far sovrapporre i due PDAG: aggiunta o cancellazione di un arco non orientato e aggiunta, rimozione o cambio di orientazione di un arco orientato. Per ogni DAG ottenuto in uscita è stata misurata la sua SHD rispetto alla struttura (nota) della rete di riferimento. Evidentemente, un algoritmo è tanto migliore nel ricostruire la struttura quanto più la SHD di tale struttura rispetto al riferimento è bassa.

- Inoltre, poichè trovare gli archi corretti di un grafo può essere visto come un problema di classificazione binario (arco presente/arco assente), abbiamo valutato **precision** e **recall**. In tale contesto, se definiamo come:

**vero positivo** (TP, true positive) un arco esistente nella rete originale e correttamente individuato dall'algoritmo in esame;

**vero negativo** (TN, false negative) un arco *non* presente nella rete originale e correttamente *non* individuato dall'algoritmo;

**falso positivo** (FP, false positive) un arco *non* esistente nella rete originale, ma che l'algoritmo in esame identifica erroneamente come esistente;

**falso negativo** (FN, true negative) un arco esistente nella rete originale ma *non* individuato,

possiamo definire:

**Precision** come la frazione di archi correttamente individuati come veri positivi rispetto al numero totale di archi indicati come esistenti dall'algoritmo. In formule:

$$\frac{TP}{TP + FP}$$

**Recall** come la frazione di archi correttamente individuati come veri positivi rispetto al numero totale di archi (esistenti o no) correttamente classificati. In formule:

$$\frac{TP}{TP + FN}$$

In altre parole, la precision esprime la probabilità che un arco individuato sia vero, mentre la recall dice quanto è probabile che un arco rilevante venga individuato.

Esiste un dilemma precision/recall, in quanto per diminuire i falsi negativi si deve spesso aumentare i falsi positivi e viceversa, per questo la valutazione delle due misure deve essere fatta in contemporanea, spesso sotto forma di curve Precision vs. Recall. Nel nostro caso, sarebbe stato complicato mostrare graficamente tutti i risultati in tale forma, così ci siamo limitati a mettere a confronto precision, recall e SHD per ogni rete, tipo e probabilità di missingness con dei grafici che variano con la taglia del dataset.

- Per valutare l'utilità di utilizzare più metodi in combinazione, l'uso di curve PvsR, o anche solo di grafici di precision e recall, si è rivelato proibitivo. Abbiamo perciò preferito valutare l'**AUC (Area Under**

**the Curve, area sotto la curva**) di grafici PvsR. Tale misura è stata valutata sul grafo somma. In pratica, una volta convertiti in CPDAG i grafi ottenuti con metodi diversi sullo stesso dataset, essi sono stati sommati. Evidentemente, in questo modo, ad ogni possibile arco verrà associata una quantità pari al numero di metodi che lo individuano (massima per l'intersezione di tutti i metodi, minima e pari a zero per gli archi non individuati da alcun metodo). Questa quantità rappresenta la fiducia (in un certo senso, la probabilità) che l'uso di più metodi in combinazione assegna alla reale presenza di tale arco. Misurando la AUC delle curve PvsR aggiungendo via via algoritmi, ci aspettiamo che tale AUC aumenti se precision e recall aumentano (ad esempio perchè due metodi sono complementari), mentre diminuisca se almeno una tra le due misure diminuisce (ad esempio perchè gli archi individuati da entrambi i metodi sono falsi positivi).

### 3.1.5 Prove

Gli esperimenti eseguiti si dividono in quattro prove:

1. Sulle quattro reti più piccole (Asia, Child, Insurance, Alarm), abbiamo testato tutti gli algoritmi e le loro varianti presentate nelle sezioni precedenti. Per ogni taglia di dataset, tipologia e probabilità di missingness sono stati calcolati i valori mediani su dieci dataset di tempo, precision, recall e SHD.
2. All'aumentare del numero di nodi della rete, abbiamo visto che i tempi di alcuni algoritmi iniziavano a diventare notevoli, senza portare a nessun particolare vantaggio rispetto agli altri. Di conseguenza, abbiamo deciso di valutare quali metodi mostravano la miglior qualità e/o il miglior rapporto tempo/qualità. Per far ciò, abbiamo calcolato le AUC delle curve PvsR utilizzando il seguente approccio greedy: per ogni rete (per ogni taglia di dataset, tipologia e probabilità di missingness sono stati calcolati i valori mediani su dieci dataset) abbiamo individuato il metodo che presentava la AUC massima; a questo metodo abbiamo sommato gli altri metodi, per vedere se la loro somma faceva aumentare la AUC; se sì, abbiamo selezionato la somma con AUC massimo e così via, finchè l'aggiunta di un metodo non provocasse un calo della AUC (nel qual caso il metodo non è stato aggiunto).
3. I metodi risultati più interessanti (SEM con BDeu a partire dalla rete MWST-EM e K-nn+MMHC) sono stati testati sulle due reti più grandi (HailFinder e Hepar2). In pratica e per completezza, dato il basso tempo computazionale, sono state effettuate tutte le prove con SEM e k-nn già indicate in precedenza.
4. Per i soli due metodi più interessanti, sono state effettuate prove esaustive su tutti i dataset di tutte le reti onde valutare le AUC ottenute con il solo SEM+T, il solo K-nn e la somma dei due. L'obiettivo era valutare se vale la pena di utilizzare la combinazione di entrambi i metodi.
5. Infine, poichè sappiamo che SEM risente positivamente dell'inizializzazione a partire da rete non vuota, abbiamo ritenuto opportuno mettere a

confronto gli ottimi risultati di SEM+T con gli analoghi risultati ottenuti da SEM inizializzato con un albero casuale (random tree, perciò indicato come SEM+RT). I due approcci sono stati confrontati su tutte le metriche, comprese le AUC.

## 3.2 Risultati

In questa sezione vengono riportati i risultati del confronto, sotto forma di tabelle e grafici. Per permettere una più facile consultazione, si è ignorata la suddivisione in prove appena descritta e si riportano insieme i risultati di tutte le reti.

**Risultati P-R-SHD** Complessivamente, come ci si può aspettare, i risultati di performance tendono a migliorare con l'aumentare della dimensione dei dataset e con il diminuire delle percentuali di incompletezza. I meccanismi di generazione della missingness, invece, non paiono avere nessun particolare effetto sulla qualità dell'apprendimento. Per la precisione, la qualità dell'apprendimento sembra leggermente minore in caso di meccanismo MCAR. Questo risultato controintuitivo è probabilmente dovuto al procedimento di generazione usato, che potrebbe potenzialmente portare ad avere più dati mancanti in caso di missingness completamente casuale.

Il numero di nodi costituisce un importante indicatore della difficoltà che si incontrerà nell'apprendimento. I risultati di Asia, che è la nostra rete più piccola, mostrano addirittura dei risultati sorprendenti: SEM (nelle sue varianti) e GES-EM su dati incompleti tendono a presentare precisione e recall maggiore rispetto ad MMHC su dati incompleti, avendo contemporaneamente una distanza di Hamming minore. Questa situazione si ripete per qualsiasi tipo di missingness e non vale soltanto in presenza di grandi percentuali di dati mancanti (nelle nostre prove, quelle indicate con p3). Nelle reti medie -cioè Child, Insurance ed Alarm- i due metodi che hanno migliori prestazioni sono k-nn e SEM nelle sue varianti. In particolare, spesso k-nn (a partire da rete vuota) mostra una recall maggiore, mentre SEM, in special modo con score BDeu e inizializzato con MWST-EM, tende ad avere una precisione anche superiore a quella di MMHC su dati completi. Queste sono però soltanto indicazioni di massima, in quanto a volte le situazioni si invertono. Per quanto concerne le SHD, esse si mantengono maggiori rispetto al caso completo, pur se molto vicine sia rispetto al riferimento, sia tra di loro. Si nota che spesso il comportamento di SEM e SEM+T in Alarm è decisamente migliore del miglior risultato per Insurance, malgrado la prima rete abbia ben dieci nodi in più rispetto alla seconda. Ciò è probabilmente da imputarsi alla struttura della particolare rete, anche se va fatto notare come il numero di parametri di Insurance sia enormemente più elevato di quello di Alarm, malgrado la minore dimensione. Precisione e recall partono da valori discretamente elevati (30-40%) anche per dataset di soli 100 elementi ed arrivano anche ad 80-100% per dataset molto grandi (5000). Anche in questo caso, Insurance mostra un andamento peggiore rispetto sia a Child che ad Alarm. Infine, per HailFinder ed Hepar2, la qualità dell'apprendimento crolla vertiginosamente. Si fa notare come ciò dipenda più dalla carenza di dati che dalla missingness: le SHD per MMHC completo oscillano spesso intorno al 100, più del numero di archi presenti nelle reti, anche per dataset di taglia 5000.

Anzi, per i dataset più piccoli la SHD dei metodi incompleti è spesso migliore rispetto ad MMHC e si dimostrano talvolta migliori anche recall e precisione. In particolare, per HailFinder, MWST-EM ha spesso la recall maggiore, laddove per ogni altra rete è tra i metodi peggiori.

Come ci si aspettava, complete case ed available case analysis hanno mostrato andamenti pessimi, ad eccezione di reti piccole e bassa missingness. MWST-EM è quasi sempre il peggiore in assoluto, come è d'altra parte logico ricordando che impara un albero e non una rete generica. Il suo utilizzo come inizializzazione di SEM, invece, porta spesso ad un miglioramento delle performance. Per gli altri metodi, MWST-EM porta invece ad un generale e notevole degradarsi delle prestazioni.

I test di Wilcoxon pairwise tra k-nn e SEM+T restituiscono un quadro senza vincitori che risente molto della specifica rete in esame. k-nn risulta spesso ottimo per Child e Insurance, mentre SEM+T vince senza difficoltà contro ogni concorrente per Alarm e HailFinder. Nel valutare quest'ultimo caso, però, occorre tener conto che i risultati sono molto bassi per tutti i metodi, verosimilmente per a causa dei dataset troppo piccoli utilizzati. Oltre tutto, HailFinder causa un comportamento strano in tutti gli algoritmi, per cui la SHD aumenta all'aumentare della taglia del dataset. Per questo, un'analisi con dataset molto più grandi (10000 o 20000 elementi) meriterebbe di essere eseguita.

Nel proseguio, i grafici relativi a queste metriche sono presentati da pagina 157 fino a pag. 181. Per le tabelle:

- da pag. 64 fino a 81 per la SHD;
- da pag. 82 fino a 99 per la precisione;
- da pag. 100 fino a 117 per la recall.

**Tempo** In linea con gli esperimenti di Tsamardinos et al., si verifica che MMHC è un metodo velocissimo, che mantiene quasi sempre i tempi di esecuzione tra i pochi decimi di secondo per reti di 8 nodi, come Asia, ed i cinque minuti per reti come Hepar2 che ne ha nove volte tanti. I risultati con k-nn sostanzialmente si sovrappongono al caso completo, quelli con l'ACA spesso sono inferiori. In tal caso, infatti, il numero di dati da analizzare è minore ed il metodo è veloce, però i risultati peggiorano molto, per cui il vantaggio reale è assente. La CCA si dimostra veloce in presenza di grandi dataset, mentre il tempo esplose su dataset piccoli, fino a superare addirittura le 12 ore da noi imposte, coerentemente con il comportamento di MMPC che prevede di testare l'indipendenza condizionata con insiemi sempre più grandi, fino ad includere l'intera CPC. Evidentemente, con pochissimi dati MMPC non riesce ad escludere l'indipendenza se non dopo molti tentativi. Meno frequentemente, anche ACA mostra lo stesso problema (in quanto essa lavora comunque su più dati rispetto alla complete case analysis). MWST-EM è un metodo il cui tempo necessario cresce lentamente con l'aumentare della taglia del dataset ed il numero di nodi, mantenendosi sempre tra la decina di minuti e le 6-7 ore. SEM è meno veloce di MMHC, ma richiede comunque solo pochi minuti per affrontare reti anche di grandi dimensioni con ottimi risultati. Va fatto notare come il tempo complessivo di SEM inizializzato con MWST-EM sia quasi esclusivamente dovuto a quest'ultimo. Infine, GES-EM è il metodo più lento che abbiamo testato: impiega qualche minuto già laddove gli altri metodi impiegano frazioni di secondo ed il tempo necessario sale notevolmente con la

dimensione del dataset ed il numero di nodi, fino a rendere impossibile rimanere nelle 12 ore già a partire da Insurance con dataset 5000 (anzi, talvolta anche per Child). Per quanto le sue performance siano buone, non sono comunque sufficienti a giustificare la preferenza rispetto a SEM o k-nn+MMHC (che hanno prestazioni migliori in tempi ridottissimi). Certamente, il fatto che GES-EM e MWST-EM siano implementati in Matlab è, in questo senso, una penalizzazione rilevante, anche se crediamo che, almeno per GES-EM, la lentezza sia ad esso intrinseca. Infatti, per quanto i nostri risultati contraddicano le conclusioni di Borchani et al., il comportamento di GES-EM è in linea con quanto indicato per GES (su cui GES-EM è basato) da Tsamardinos et al. nella loro comparazione, ed in parte già individuato da Chickering stesso. Va d'altronde fatto notare come Borchani et al. abbiano probabilmente testato il loro metodo contro l'implementazione di SEM presente nel pacchetto BNT-SLP, certamente più lenta rispetto al LibB da noi utilizzato.

I grafici relativi al tempo sono presentati da pagina 183 fino a pag. 190, le tabelle da pag. 118 a pag. 135.

**AUC** Sulle quattro reti più piccole non si registra una combinazione di metodi che si comporti decisamente meglio rispetto alle altre. Si rilevano infatti 76 combinazioni diverse (senza tener conto dell'ordine) sulle 180 migliori ottenute (4 reti x 5 taglie x 3 tipi di missingness x 3 probabilità di missingness). Oltre tutto, spesso la AUC migliore si ottiene con un solo metodo e, anche qualora l'aggiunta di un secondo migliori i risultati, i miglioramenti sono quasi sempre trascurabili. In ogni caso, è giusto rimarcare che un approccio greedy può portare a risultati subottimi.

Un'osservazione più attenta ha mostrato che le 3 combinazioni più frequenti -circa 25 ricorrenze a testa- utilizzano erano: SEM+T; k-nn; SEM+T e k-nn. Per quanto questi risultati siano molto influenzati dalla rete in esame (ad esempio, Alarm tende a preferire SEM+T, altre reti preferiscono k-nn) si è perciò deciso di valutare estensivamente le AUC delle curve PvsR rispetto a tutti i dataset per queste combinazioni. In questo caso, si rileva che l'uso congiunto dei due metodi raramente peggiora i risultati rispetto al metodo peggiore e, nel caso, comunque in maniera trascurabile. Molto più spesso, la AUC dell'ensemble learning si aggancia al metodo migliore e mostra un miglioramento interessante rispetto all'algoritmo peggiore, pur se inferiore rispetto al migliore. D'altra parte, è difficile predire quando uno dei due algoritmi sarà migliore dell'altro, ed in questo senso può risultare utile sapere che l'uso della loro combinazione non peggiora mai risultati, ma può migliorarli rispetto al caso peggiore. Talvolta, anzi, l'uso congiunto migliora le AUC anche oltre i risultati singoli.

Anche in questo caso, i test statistici non mostrano vincitori netti, ma si può anzi dire che ogni rete è un caso a sè stante.

I grafici relativi alle AUC con metodo greedy sono presentati da pagina 194 fino a pag. 199. Per l'ensemble learning relativo ai soli k-nn e SEM+T, i grafici vanno da pag. 203 a pag. 209, le tabelle da pag. 136 a pag. 153.

**Inizializzazione con albero casuale** I confronti tra SEM+T e SEM+RT non mostrano quasi mai variazioni significative. In genere, SEM+RT è leggermente peggiore di SEM+T ma, alla luce dei test di Wilcoxon, le variazioni sono rilevanti esclusivamente per Alarm, rete per cui SEM+T si dimostra sempre il metodo

migliore. Altre variazioni significative frequenti si registrano solo per HailFinder. In questo caso, però, è legittimo un dubbio legato al numero di nodi e di parametri della rete e alla conseguente dimensione insufficiente anche dei dataset di taglia 5000. Occorrerebbe pertanto eseguire prove con dataset più estesi al fine di valutare la reale efficacia di SEM+T rispetto a SEM+RT, in quanto, in ogni caso, le prestazioni sono estremamente basse per entrambi i metodi.

Anche per quanto riguarda le AUC, SEM+T mostra variazioni rilevanti e significative per le stesse reti, mentre in altri casi è SEM+RT (o k-nn, in senso complessivo) a fornire i risultati migliori. A meno di trovare un criterio per individuare le reti che traggono benefici da MWST-EM, l'aumento di qualità di SEM+T rispetto a SEM+RT è spesso troppo fiavole per giustificare il costo in termini di tempo (sia di esecuzione che implementativo).

I risultati relativi a SEM+RT sono in coda alle tabelle delle varie metriche già menzionate. I grafici di confronto tra SEM+T e SEM+RT vanno da pag. 210 a pag. 236.

### 3.2.1 Tabelle

Le tabelle qui riportate mostrano nella direzione orizzontale la variazione della taglia del dataset, mentre in quella verticale i metodi e le diverse probabilità di missingness. Ogni tabella mostra i risultati mediani di una misura su dieci dataset, per rete e tipo di missingness. Nella prima riga è indicato il risultato mediano ottenuto con MMHC applicato sui dataset completi, così da fornire un riferimento. Tra parentesi è riportato il numero di dataset che hanno effettivamente riportato risultati validi nell'arco delle 12 ore. Per ogni taglia, tipo e probabilità di missingness in grassetto è individuata la misura migliore (massima per precision e recall, minima per SHD e tempo). Si sono esclusi dal confronto MMHC, in quanto ha solo funzione di riferimento, e SEM+RT, che ha solo funzione di controllo per SEM+T.

Per SHD, precisione, recall, tempo, sono stati eseguiti test statistici pairwise (Wilcoxon signed rank sum test, livello di significatività  $\alpha = 0.05$ ) tra k-nn e SEM+T (evidenziate nelle tabelle con il simbolo \*) e tra SEM+T e SEM+RT (nelle tabelle con il simbolo †). Per il confronto complessivo tra le AUC, si è utilizzato un test Kruskal-Wallis con correzione di Bonferroni tra k-nn, SEM+T, somma k-nn e SEM+T (livello di significatività  $\alpha = 0.05/3 = 0.0167$ ). Qualora esso abbia dato risultato positivo, sono stati eseguiti test pairwise (Wilcoxon signed rank sum test, livello di significatività  $\alpha = 0.05$ ) tra tutte le coppie possibili dei tre metodi e si è evidenziato con un asterisco (\*) un valore soltanto quando esso sia risultato statisticamente maggiore rispetto ad entrambi gli altri.

Si mette in evidenza come SEM (o, meglio, LibB) mostri uno strano comportamento per i dataset di dimensione 5000 e probabilità di missingness elevata (p3). Infatti, spesso non ci sono i risultati, poichè in tal caso LibB esce inaspettatamente dal programma, senza apparentemente segnalare errori. Verosimilmente, si tratta di un limite implementativo, più che algoritmico.

## Structural Hamming Distance

SHD Asia mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	6.5(10)	6.5(10)	6.0(10)	6.0(10)	5.0(10)
	1	11.0(10)	10.0(10)	10.0(10)	10.0(10)	10.0(10)
	3	10.0(10)	10.0(10)	9.5(10)	10.0(10)	10.0(10)
MWST-EM	1	7.5(10)	7.0(10)	6.5(10)	6.0(10)	5.5(10)
	2	9.0(10)	8.5(10)	6.0(10)	6.5(10)	5.5(10)
	3	11.5(10)	9.0(10)	8.0(10)	7.0(10)	6.5(10)
CC	1	7.0(10)	7.5(10)	7.5(10)	8.0(10)	8.5(10)
	2	8.0(10)	7.5(10)	7.0(10)	7.0(10)	8.0(10)
	3	8.0(10)	7.0(10)	7.5(10)	7.0(10)	8.5(10)
AC	1	7.5(10)	9.0(10)	8.0(10)	8.0(10)	5.0(10)
	2	7.5(10)	9.0(10)	9.0(10)	9.0(10)	8.0(10)
	3	8.0(10)	8.0(10)	7.0(10)	10.0(10)	8.0(10)
AC+T	1	11.0(10)	10.0(10)	10.0(10)	11.0(10)	11.0(10)
	2	11.0(10)	10.5(10)	10.0(10)	10.5(10)	10.5(10)
	3	11.0(10)	10.5(10)	10.0(10)	10.0(10)	10.0(10)
k10	1	6.5(10)	6.5(10)	7.0(10)	6.0(10)	4.0(10)
	2	9.0(10)	8.5(10)	7.0(10)	6.0(10)	4.0(10)
	3	9.0(10)	8.5(10)	7.5(10)	7.5(10)	5.5(10)
k10+T	1	8.0(10)	7.0(10)	7.5(10)	9.0(10)	9.0(10)
	2	7.5(10)	8.0(10)	8.0(10)	9.0(10)	9.0(10)
	3	8.0(10)	8.0(10)	7.5(10)	9.0(10)	9.5(10)
SEM	1	<b>6.0(10)</b>	5.0(10)	3.0(10)	2.5(10)	2.0(10)
	2	7.5(10)	<b>4.5(10)</b>	4.0(10)	3.5(10)	<b>2.0(10)</b>
	3	7.0(10)	8.0(10)	<b>4.0(10)</b>	<b>3.0(10)</b>	<b>2.0(10)</b>
SEM+T	1	6.5(10)	5.0(10)	2.5(10)*	<b>2.0(10)*</b>	2.0(10)
	2	6.5(10)*	4.5(10)*	<b>3.5(10)*</b>	<b>2.0(10)*</b>	2.0(10)
	3	7.0(10)*	<b>6.5(10)*</b>	4.0(10)*	4.0(10)*	2.0(10)*
SEM.bic	1	7.0(10)	5.0(10)	2.5(10)	2.0(10)	2.0(10)
	2	7.5(10)	5.0(10)	4.5(10)	3.5(10)	2.0(10)
	3	7.5(10)	7.0(10)	6.0(10)	4.5(10)	2.0(10)
SEM.bic+T	1	6.5(10)	<b>4.0(10)</b>	<b>2.0(10)</b>	2.0(10)	2.0(10)
	2	7.0(10)	4.5(10)	3.5(10)	4.0(10)	2.0(10)
	3	6.5(10)	7.0(10)	5.5(10)	4.5(10)	2.0(10)
GES-EM	1	6.0(10)	5.0(10)	3.0(10)	2.0(10)	<b>1.5(10)</b>
	2	7.0(10)	6.0(10)	4.0(10)	2.5(10)	2.0(10)
	3	<b>6.0(10)</b>	8.5(10)	6.0(10)	6.0(10)	2.0(10)
GES-EM+T	1	6.5(10)	6.5(10)	6.5(10)	7.5(10)	8.5(10)
	2	<b>6.0(10)</b>	5.5(10)	5.5(10)	4.5(10)	5.0(10)
	3	7.0(10)	6.5(10)	7.0(10)	8.0(10)	5.5(10)
SEM+RT	1	6.0(10)	5.0(10)	3.0(10)	3.5(10)	3.0(10)
	2	7.5(10)	5.0(10)	3.0(10)	3.5(10)	2.0(10)
	3	8.0(10)	7.5(10)	4.0(10)	5.0(10)	2.0(10)

SHD Asia mcar SHD incompleti

SHD Child mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	20.0(10)	13.0(10)	9.0(10)	6.0(10)	3.5(10)
MWST-EM	1	43.0(10)	42.0(10)	43.0(10)	43.0(10)	42.0(10)
	2	42.0(10)	43.0(10)	43.0(10)	43.0(10)	42.0(10)
	3	42.0(10)	43.0(10)	43.0(10)	42.0(10)	40.0(10)
CC	1	30.5(10)	19.0(10)	15.5(10)	9.5(10)	5.5(10)
	2	73.5(10)	68.0(10)	38.0(10)	25.5(10)	12.0(10)
	3	69.0(3)	70.5(4)	76.0(7)	94.0(10)	38.5(10)
CC+T	1	27.0(10)	27.0(10)	28.0(10)	29.0(10)	35.0(10)
	2	32.5(10)	32.5(10)	27.5(10)	28.0(10)	28.0(10)
	3	41.0(3)	41.0(4)	36.0(7)	33.5(10)	27.5(10)
AC	1	29.0(10)	18.0(10)	13.5(10)	8.0(10)	10.0(10)
	2	42.0(10)	33.0(10)	18.5(10)	18.5(10)	13.0(10)
	3	48.0(10)	35.5(10)	28.0(10)	27.0(10)	17.0(10)
AC+T	1	38.0(10)	38.0(10)	42.0(10)	42.0(10)	41.0(10)
	2	42.0(10)	43.0(10)	43.0(10)	43.0(10)	41.0(10)
	3	42.0(10)	43.0(10)	43.0(10)	42.0(10)	39.0(10)
k10	1	19.5(10)	15.0(10)	11.0(10)	<b>7.0(10)*</b>	<b>2.5(10)</b>
	2	22.0(10)	16.5(10)*	<b>11.0(10)*</b>	<b>9.0(10)*</b>	<b>2.5(10)*</b>
	3	<b>23.0(10)</b>	<b>17.5(10)*</b>	<b>13.5(10)*</b>	<b>11.0(10)*</b>	5.0(10)*
k10+T	1	28.0(10)	28.5(10)	30.0(10)	32.0(10)	37.5(10)
	2	27.0(10)	29.0(10)	31.0(10)	33.0(10)	37.0(10)
	3	28.0(10)	28.0(10)	31.0(10)	33.5(10)	37.0(10)
SEM	1	19.0(10)	15.0(10)	12.0(10)	7.5(10)	3.0(10)
	2	23.0(10)	<b>16.0(10)</b>	13.5(10)	12.5(10)	4.0(10)
	3	24.0(10)	22.0(10)	17.5(10)	15.5(10)	5.0(10)
SEM+T	1	18.5(10)	17.5(10)	12.5(10)	8.5(10)	3.5(10)
	2	23.0(10)	20.5(10)	17.0(10)	11.5(10)	10.0(10)
	3	25.0(10)	23.0(10)	20.5(10)	18.0(10)	11.0(9)
SEM.bic	1	<b>18.0(10)</b>	17.5(10)	11.5(10)	8.5(10)	3.0(10)
	2	<b>21.0(10)</b>	16.0(10)	11.5(10)	11.5(10)	3.0(10)
	3	24.5(10)	19.0(10)	16.0(10)	13.0(10)	<b>4.5(10)</b>
SEM.bic+T	1	19.0(10)	17.5(10)	14.0(10)	8.0(10)	5.0(10)
	2	23.0(10)	20.0(10)	16.5(10)	11.5(10)	6.0(9)
	3	25.0(10)	22.0(10)	23.0(10)	16.5(10)	8.0(9)
GES-EM	1	18.5(10)	<b>14.0(10)</b>	<b>10.0(10)</b>	8.0(10)	4.5(10)
	2	22.0(10)	17.0(10)	12.0(10)	10.5(10)	7.0(9)
	3	23.5(10)	19.5(10)	15.0(10)	13.0(10)	8.5(10)
GES-EM+T	1	27.5(10)	25.0(9)	28.0(10)	24.5(10)	20.5(10)
	2	28.0(10)	30.0(10)	30.5(10)	26.5(6)	20.5(8)
	3	-	-	-	-	23.0(3)
SEM+RT	1	20.0(10)	17.5(10)	12.5(10)	8.0(10)	6.0(10)
	2	21.5(10)	18.0(10)†	16.5(10)	13.5(10)	5.5(10)†
	3	24.0(10)	21.0(10)†	21.0(10)	20.0(10)	9.0(10)

SHD Child mcar SHD incompleti

SHD Insurance mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	42.5(10)	37.5(10)	31.5(10)	31.5(10)	21.0(10)
MWST-EM	1	64.0(10)	63.0(10)	63.5(10)	61.5(10)	58.5(10)
	2	64.0(10)	61.0(10)	62.0(10)	62.0(10)	58.5(10)
	3	61.5(10)	62.0(10)	61.0(10)	61.0(10)	58.5(10)
CC	1	60.0(10)	53.5(10)	37.0(10)	37.0(10)	26.0(10)
	2	154.0(5)	168.0(7)	129.0(10)	103.5(10)	49.5(10)
	3	-	129.0(3)	127.0(1)	138.0(1)	160.0(9)
CC+T	1	51.5(10)	51.0(10)	52.0(10)	52.0(10)	53.0(10)
	2	61.0(6)	55.5(10)	53.5(10)	54.0(10)	51.0(10)
	3	-	63.0(3)	66.0(1)	61.0(1)	57.5(10)
AC	1	62.5(10)	50.0(10)	43.0(10)	44.0(10)	42.0(10)
	2	103.0(10)	76.5(10)	53.0(10)	50.5(10)	41.0(10)
	3	121.5(10)	88.0(10)	63.5(10)	56.5(10)	42.0(10)
AC+T	1	57.0(10)	60.0(10)	60.0(10)	61.0(10)	60.5(10)
	2	63.5(10)	61.0(10)	63.0(10)	63.5(10)	62.0(10)
	3	61.0(10)	62.0(10)	61.0(10)	61.5(10)	60.0(10)
k10	1	46.5(10)	<b>41.5(10)</b>	<b>33.5(10)*</b>	<b>29.5(10)*</b>	<b>20.0(10)*</b>
	2	52.0(10)	<b>42.0(10)*</b>	<b>38.5(10)</b>	<b>32.5(10)*</b>	<b>25.5(10)</b>
	3	56.5(10)	48.0(10)	<b>45.0(10)</b>	<b>38.5(10)*</b>	<b>32.0(10)</b>
k10+T	1	51.5(10)	52.0(10)	53.0(10)	52.5(10)	54.5(10)
	2	52.0(10)	51.5(10)	52.0(10)	53.5(10)	54.5(10)
	3	52.0(10)	52.5(10)	52.5(10)	52.5(10)	54.5(10)
SEM	1	46.5(10)	42.5(10)	38.5(10)	37.5(10)	25.5(10)
	2	<b>49.5(10)</b>	47.5(10)	42.5(10)	40.0(10)	29.0(9)
	3	50.5(10)	48.0(10)	45.5(10)	41.0(10)	38.5(4)
SEM+T	1	46.0(10)†	44.0(10)	42.5(10)	38.5(10)	25.5(10)
	2	50.0(10)†	48.0(10)	44.5(10)	44.5(10)	32.0(7)
	3	51.0(10)*	50.0(10)	48.0(10)	48.0(10)	41.0(1)
SEM.bic	1	48.5(10)	44.0(10)	42.5(10)	42.0(10)	31.0(10)
	2	52.0(10)	47.5(10)	44.5(10)	42.5(10)	32.0(9)
	3	50.5(10)	49.0(10)	46.0(10)	43.5(10)	41.5(6)
SEM.bic+T	1	<b>45.5(10)</b>	44.5(10)	44.0(10)	40.5(10)	26.0(10)
	2	50.0(10)	47.5(10)	48.0(10)	45.5(10)	33.0(8)
	3	<b>49.5(10)</b>	51.0(10)	48.0(10)	47.5(10)	45.0(2)
GES-EM	1	48.0(10)	43.5(10)	38.0(9)	37.0(9)	-
	2	50.0(10)	45.5(10)	42.0(9)	42.0(9)	-
	3	51.0(9)	<b>47.5(10)</b>	45.5(10)	46.5(10)	-
GES-EM+T	1	52.0(10)	54.5(10)	52.5(10)	53.5(10)	-
	2	50.0(10)	50.0(10)	48.0(10)	48.5(10)	-
	3	50.5(10)	51.0(10)	57.0(10)	56.0(8)	-
SEM+RT	1	49.0(10)	45.5(10)	40.0(10)	37.0(10)	25.0(10)
	2	52.5(10)	48.0(10)	45.0(10)	43.5(10)	36.0(7)
	3	51.5(10)	49.0(10)	49.5(10)	52.0(9)	45.5(2)

SHD Insurance mcar SHD incompleti

SHD Alarm1 mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	41.5(10)	34.0(10)	20.5(10)	20.0(10)	14.0(10)
MWST-EM	1	56.5(10)	55.0(10)	56.0(10)	57.0(10)	60.0(10)
	2	56.5(10)	56.0(10)	58.5(10)	55.0(10)	65.0(10)
	3	60.0(10)	56.5(10)	56.5(10)	57.0(10)	57.5(10)
CC	1	197.5(10)	117.5(10)	50.0(10)	31.0(10)	23.0(10)
	2	134.0(3)	133.0(4)	200.0(2)	129.0(5)	170.0(10)
	3	-	-	74.0(1)	118.0(1)	137.0(1)
CC+T	1	49.0(10)	49.0(10)	45.0(10)	45.5(10)	48.5(10)
	2	55.5(4)	56.5(6)	54.0(3)	55.0(10)	51.0(10)
	3	-	-	57.0(1)	57.0(1)	58.0(1)
AC	1	73.5(10)	59.0(10)	50.0(10)	40.5(10)	33.5(10)
	2	136.0(10)	76.0(10)	64.0(10)	48.5(10)	39.0(10)
	3	171.0(10)	91.5(10)	63.0(10)	54.5(10)	40.5(10)
AC+T	1	54.0(10)	52.5(10)	55.0(10)	56.5(10)	58.5(10)
	2	56.5(10)	56.0(10)	58.5(10)	56.0(10)	59.5(10)
	3	58.5(10)	57.0(10)	56.5(10)	56.5(10)	59.0(10)
k10	1	43.0(10)	36.0(10)	25.5(10)	22.5(10)	19.5(10)
	2	51.0(10)	40.5(10)	31.5(10)	29.0(10)	25.0(10)*
	3	62.0(10)	52.5(10)	41.5(10)	35.5(10)	37.0(10)
k10+T	1	45.0(10)	44.5(10)	43.5(10)	45.0(10)	53.5(10)
	2	45.0(10)	45.5(10)	45.0(10)	44.0(10)	56.0(10)
	3	46.0(10)	45.5(10)	45.5(10)	44.5(10)	51.5(10)
SEM	1	33.5(10)	31.5(10)	23.0(10)	21.5(10)	25.0(10)
	2	39.5(10)	36.5(10)	30.0(10)	29.0(9)	<b>19.5(8)</b>
	3	48.5(10)	<b>41.5(10)</b>	42.5(10)	29.5(8)	31.0(1)
SEM+T	1	<b>23.5(10)*†</b>	<b>22.5(10)*†</b>	<b>14.5(10)*†</b>	<b>13.0(10)*†</b>	<b>8.5(10)†</b>
	2	<b>39.0(10)*</b>	<b>29.0(10)*†</b>	<b>26.5(10)*†</b>	<b>21.0(10)*†</b>	36.0(8)
	3	<b>44.5(10)*†</b>	45.5(10)*	<b>33.5(10)*†</b>	<b>28.0(9)*†</b>	<b>23.0(5)</b>
SEM.bic	1	37.5(10)	41.0(10)	28.0(10)	30.5(10)	25.0(10)
	2	42.5(10)	41.5(10)	32.5(10)	33.0(10)	26.0(7)
	3	47.0(10)	45.0(10)	43.5(10)	29.5(10)	46.0(3)
SEM.bic+T	1	40.0(10)	36.0(10)	27.0(10)	22.5(10)	19.5(10)
	2	46.0(10)	39.5(10)	31.5(10)	28.5(10)	32.0(10)
	3	48.5(10)	46.0(10)	37.5(10)	34.0(9)	27.5(6)
GES-EM	1	39.0(10)	38.0(9)	31.5(10)	-	-
	2	42.5(10)	37.0(10)	35.0(10)	33.5(6)	-
	3	46.0(10)	44.5(10)	40.0(10)	34.5(10)	-
GES-EM+T	1	53.5(10)	54.0(10)	45.0(10)	46.0(7)	-
	2	53.0(10)	45.0(10)	-	39.0(3)	-
	3	59.0(10)	54.5(10)	48.0(9)	-	-
SEM+RT	1	35.0(10)	33.5(10)	33.0(10)	22.5(10)	29.0(9)
	2	38.0(10)	41.5(10)	32.0(10)	36.0(9)	43.0(5)
	3	50.5(10)	46.5(10)	50.0(10)	40.0(7)	-

SHD Alarm1 mcar SHD incompleti

SHD HailFinder mcar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	95.5(10)	100.5(10)	105.0(10)	103.0(10)	101.5(10)
MWST-EM	1	109.5(10)	109.0(10)	108.0(10)	106.5(10)	101.0(10)
	2	111.0(10)	110.0(10)	109.0(10)	106.0(10)	104.5(10)
	3	109.0(10)	107.5(10)	107.5(10)	105.0(10)	102.0(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	90.0(10)	97.0(10)	108.0(10)	107.0(10)	108.0(10)
	2	90.0(10)	95.0(10)	107.0(10)	108.5(10)	113.0(10)
	3	90.0(10)	93.0(10)	104.5(10)	109.0(10)	115.5(10)
k10+T	1	<b>76.5(10)</b>	<b>79.0(10)</b>	<b>80.0(10)</b>	<b>81.5(10)</b>	<b>87.0(10)</b>
	2	<b>77.0(10)</b>	<b>79.0(10)</b>	<b>81.0(10)</b>	<b>83.0(10)</b>	<b>85.5(10)</b>
	3	75.5(10)	<b>78.0(10)</b>	<b>81.0(10)</b>	<b>82.5(10)</b>	<b>84.0(10)</b>
SEM	1	84.0(10)	89.5(10)	101.0(10)	109.0(10)	107.5(8)
	2	79.5(10)	84.5(10)	98.5(10)	109.5(8)	116.0(1)
	3	<b>70.0(10)</b>	81.0(10)	90.5(10)	103.0(9)	-
SEM+T	1	86.5(10)*	92.0(10)*	102.5(10)*	110.0(10)	111.0(8)
	2	82.0(10)*	86.0(10)*	100.5(10)*	111.5(8)	114.5(2)
	3	77.5(10)*	84.5(10)*	95.5(10)*	106.5(10)	-
SEM.bic	1	83.5(10)	91.0(10)	105.5(10)	110.0(10)	109.0(5)
	2	80.0(10)	86.0(10)	101.0(10)	110.0(6)	-
	3	72.0(10)	82.5(10)	93.0(10)	107.0(7)	-
SEM.bic+T	1	85.0(10)	92.0(10)	107.0(10)	107.5(10)	112.5(6)
	2	83.0(10)	88.0(10)	103.5(10)	112.0(7)	-
	3	80.5(10)	88.0(10)	96.0(10)	108.0(9)	-
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	86.0(10)	90.5(10)	102.0(10)	109.0(10)	112.5(8)
	2	81.0(10)†	85.0(10)	100.0(10)	110.0(8)	-
	3	72.0(10)†	82.5(10)†	92.0(9)†	105.0(8)	-

SHD HailFinder mcar SHD incompleti

SHD Hepar2 mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	171.5(10)	151.0(10)	121.0(10)	104.0(10)	70.5(10)
MWST-EM	1	176.0(10)	171.0(10)	173.5(10)	172.5(10)	174.5(10)
	2	173.5(10)	174.5(10)	176.0(10)	174.0(10)	173.5(10)
	3	173.5(10)	175.0(10)	175.5(10)	172.5(10)	175.0(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	175.5(10)	153.5(10)	122.0(10)	108.0(10)	71.5(10)
	2	184.0(10)	157.0(10)	135.0(10)	118.0(10)	82.0(10)
	3	187.5(10)	167.5(10)	142.0(10)	130.5(10)	88.0(10)
k10+T	1	126.0(10)	124.5(10)	123.5(10)	124.5(10)	126.0(10)
	2	128.0(10)	125.0(10)	124.0(10)	124.5(10)	124.5(10)
	3	126.0(10)	125.0(10)	125.5(10)	123.5(10)	125.0(10)
SEM	1	<b>118.5(10)</b>	114.5(10)	104.5(10)	<b>91.5(10)</b>	67.5(10)
	2	121.0(10)	117.0(10)	<b>109.5(10)</b>	<b>97.0(10)</b>	<b>72.5(10)</b>
	3	122.0(10)	120.0(10)	115.5(10)	<b>105.0(10)</b>	<b>80.0(10)</b>
SEM+T	1	119.0(10)*	<b>113.5(10)*</b>	<b>102.0(10)*</b>	93.5(10)*	77.0(10)
	2	<b>120.5(10)*</b>	<b>115.5(10)*</b>	111.0(10)*	98.5(10)*	79.5(10)
	3	121.0(10)*	<b>118.5(10)*†</b>	<b>114.5(10)*</b>	105.5(10)*	85.0(10)
SEM.bic	1	119.0(10)	114.5(10)	103.0(10)	93.5(10)	<b>67.0(10)</b>
	2	121.5(10)	119.0(10)	110.0(10)	98.0(10)	73.5(10)
	3	122.0(10)	120.0(10)	115.5(10)	108.0(10)	80.5(10)
SEM.bic+T	1	119.0(10)	116.0(10)	105.0(10)	95.5(10)	79.5(10)
	2	122.0(10)	118.0(10)	110.0(10)	100.5(10)	81.5(10)
	3	<b>120.0(10)</b>	119.5(10)	116.0(10)	107.0(10)	86.5(10)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	118.5(10)	116.0(10)	102.5(10)	95.0(10)	68.5(10)†
	2	121.5(10)	116.5(10)	111.0(10)	98.5(10)	72.5(10)†
	3	121.5(10)	120.0(10)	115.5(10)	105.5(10)	80.0(10)

SHD Hepar2 mcar SHD incompleti

SHD Asia mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	6.5(10)	6.5(10)	6.0(10)	6.0(10)	5.0(10)
MWST-EM	1	9.0(10)	10.5(10)	10.0(10)	10.5(10)	11.0(10)
	2	10.0(10)	11.0(10)	10.0(10)	10.0(10)	11.0(10)
	3	11.0(10)	10.0(10)	10.0(10)	10.0(10)	11.5(10)
CC	1	8.0(10)	6.5(10)	6.0(10)	6.0(10)	5.0(10)
	2	8.5(10)	6.5(10)	6.5(10)	7.0(10)	5.0(10)
	3	10.5(10)	9.0(10)	6.0(10)	6.0(10)	6.5(10)
CC+T	1	7.5(10)	7.0(10)	8.0(10)	8.0(10)	9.0(10)
	2	7.5(10)	8.0(10)	7.0(10)	7.5(10)	9.0(10)
	3	7.5(10)	7.5(10)	8.0(10)	7.0(10)	9.0(10)
AC	1	7.5(10)	8.5(10)	7.0(10)	7.0(10)	7.5(10)
	2	7.5(10)	8.5(10)	7.0(10)	8.0(10)	7.0(10)
	3	7.5(10)	8.0(10)	7.0(10)	8.0(10)	8.0(10)
AC+T	1	9.0(10)	9.0(10)	10.0(10)	10.0(10)	10.5(10)
	2	8.0(10)	9.0(10)	10.0(10)	11.0(10)	10.5(10)
	3	8.5(10)	9.0(10)	10.0(10)	10.0(10)	11.0(10)
k10	1	7.0(10)	7.0(10)	6.5(10)	5.5(10)	5.5(10)
	2	7.0(10)	6.5(10)	7.0(10)	6.0(10)	4.5(10)
	3	7.0(10)	7.0(10)	6.5(10)	8.0(10)	7.5(10)
k10+T	1	7.0(10)	8.0(10)	9.0(10)	8.5(10)	9.0(10)
	2	7.0(10)	8.0(10)	8.5(10)	9.0(10)	10.0(10)
	3	7.5(10)	7.5(10)	8.0(10)	8.5(10)	10.0(10)
SEM	1	6.0(10)	<b>4.5(10)</b>	3.5(10)	<b>2.0(10)</b>	2.0(10)
	2	7.0(10)	<b>5.0(10)</b>	4.0(10)	2.5(10)	3.0(10)
	3	7.0(10)	6.5(10)	5.5(10)	<b>3.5(10)</b>	3.5(10)
SEM+T	1	<b>5.5(10)*</b>	5.0(10)*	<b>2.5(10)*</b>	2.5(10)*	2.0(10)
	2	<b>5.5(10)</b>	5.0(10)	<b>2.5(10)*</b>	3.0(10)*	<b>2.0(10)</b>
	3	7.5(10)	7.0(10)	3.5(10)	4.0(10)*	3.0(10)*
SEM.bic	1	6.5(10)	6.0(10)	3.0(10)	2.0(10)	5.0(10)
	2	8.0(10)	5.5(10)	4.0(10)	<b>2.0(10)</b>	2.0(10)
	3	7.5(10)	7.0(10)	5.0(10)	3.5(10)	3.0(10)
SEM.bic+T	1	6.0(10)	5.0(10)	2.5(10)	2.0(10)	2.0(10)
	2	6.0(10)	5.0(10)	3.0(10)	3.0(10)	2.0(10)
	3	7.0(10)	<b>5.5(10)</b>	<b>3.0(10)</b>	4.0(10)	<b>2.0(10)</b>
GES-EM	1	6.0(10)	6.5(10)	3.5(10)	3.0(10)	<b>1.5(10)</b>
	2	6.5(10)	6.0(10)	4.0(9)	2.5(10)	2.0(9)
	3	7.0(10)	6.5(10)	5.0(10)	6.0(10)	3.0(10)
GES-EM+T	1	7.0(9)	7.0(10)	8.0(10)	7.0(10)	11.0(10)
	2	6.5(10)	6.0(10)	7.0(10)	7.0(10)	8.0(9)
	3	<b>6.0(9)</b>	7.0(10)	7.5(10)	8.0(10)	10.0(10)
SEM+RT	1	5.0(10)	5.5(10)	2.0(10)	3.0(10)	2.0(10)
	2	5.5(10)	4.5(10)	4.5(10)	2.5(10)	2.0(10)
	3	7.5(10)	8.0(10)	3.5(10)	5.5(10)	5.0(10)

SHD Asia mar SHD incompleti

SHD Child mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	20.0(10)	13.0(10)	9.0(10)	6.0(10)	3.5(10)
MWST-EM	1	42.5(10)	43.0(10)	43.0(10)	42.5(10)	42.0(10)
	2	42.5(10)	43.0(10)	43.0(10)	41.5(10)	40.0(10)
	3	42.5(10)	42.0(10)	42.5(10)	42.5(10)	42.0(10)
CC	1	41.0(10)	28.0(10)	17.5(10)	11.0(10)	7.0(10)
	2	79.0(10)	40.5(10)	28.5(10)	19.0(10)	10.5(10)
	3	96.0(9)	72.0(10)	57.0(10)	35.0(10)	17.0(10)
CC+T	1	29.0(10)	27.5(10)	28.0(10)	28.5(10)	31.5(10)
	2	33.5(10)	29.0(10)	29.0(10)	27.5(10)	29.0(10)
	3	36.0(9)	30.5(10)	29.5(10)	28.0(10)	28.0(10)
AC	1	38.0(10)	24.0(10)	20.5(10)	17.5(10)	14.0(10)
	2	43.0(10)	27.0(10)	24.0(10)	21.0(10)	15.5(10)
	3	45.0(10)	34.0(10)	25.0(10)	21.0(10)	16.5(10)
AC+T	1	36.0(10)	37.5(10)	38.0(10)	37.0(10)	39.5(10)
	2	39.0(10)	38.0(10)	38.5(10)	39.0(10)	38.5(10)
	3	37.5(10)	38.0(10)	38.0(10)	39.0(10)	39.5(10)
k10	1	20.5(10)	<b>14.5(10)</b>	<b>10.0(10)</b>	<b>8.0(10)</b>	<b>2.5(10)*</b>
	2	20.5(10)	<b>14.5(10)*</b>	13.0(10)*	10.0(10)	<b>3.0(10)*</b>
	3	26.0(10)	<b>15.5(10)*</b>	<b>13.5(10)</b>	<b>10.0(10)*</b>	<b>5.5(10)*</b>
k10+T	1	27.5(10)	29.0(10)	31.0(10)	32.0(10)	37.0(10)
	2	28.0(10)	28.0(10)	31.5(10)	33.0(10)	36.5(10)
	3	29.0(10)	29.0(10)	31.0(10)	33.0(10)	38.0(10)
SEM	1	20.5(10)	14.5(10)	15.0(10)	8.0(10)	5.0(10)
	2	22.5(10)	16.5(10)	17.5(10)	10.5(10)	7.0(10)
	3	23.0(10)	18.0(10)	19.0(10)	14.5(10)	8.5(10)
SEM+T	1	21.0(10)	15.5(10)	14.5(10)	11.0(10)	4.0(10)
	2	<b>20.0(10)</b>	18.5(10)†	16.5(10)	11.5(10)	8.5(10)
	3	22.0(10)	19.0(10)	17.0(10)	16.0(10)	12.0(10)
SEM.bic	1	21.5(10)	16.0(10)	12.5(10)	9.0(10)	5.0(10)
	2	22.5(10)	18.0(10)	16.0(10)	<b>9.5(10)</b>	5.0(10)
	3	22.0(10)	21.0(10)	16.0(10)	11.5(10)	7.0(10)
SEM.bic+T	1	22.0(10)	18.0(10)	11.5(10)	13.5(10)	5.0(10)
	2	23.0(10)	18.0(10)	16.0(10)	13.5(10)	6.0(10)
	3	23.0(10)	21.0(10)	17.5(10)	15.5(10)	9.0(10)
GES-EM	1	<b>19.5(10)</b>	14.5(10)	10.0(10)	8.0(10)	9.0(8)
	2	21.5(10)	16.5(10)	<b>12.0(10)</b>	10.0(10)	7.5(10)
	3	<b>21.5(10)</b>	16.5(10)	14.0(10)	11.5(10)	9.0(9)
GES-EM+T	1	27.0(10)	27.5(8)	28.0(9)	28.0(10)	23.0(9)
	2	28.0(10)	28.0(10)	28.5(10)	27.0(7)	23.5(6)
	3	-	-	-	-	31.0(7)
SEM+RT	1	21.5(10)	17.5(10)	14.0(10)	12.0(10)	5.0(10)
	2	22.0(10)	18.5(10)	17.5(10)	12.0(10)	6.0(10)
	3	23.5(10)	18.0(10)	19.0(10)	14.0(10)	11.0(10)

SHD Child mar SHD incompleti

SHD Insurance mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	42.5(10)	37.5(10)	31.5(10)	31.5(10)	21.0(10)
MWST-EM	1	62.5(10)	62.0(10)	62.5(10)	61.5(10)	58.5(10)
	2	64.0(10)	62.5(10)	61.5(10)	63.0(10)	60.0(10)
	3	63.0(10)	60.5(10)	61.0(10)	60.5(10)	59.5(10)
CC	1	86.5(10)	64.5(10)	51.0(10)	42.0(10)	33.0(10)
	2	134.5(4)	89.0(9)	69.5(10)	56.0(9)	40.0(10)
	3	164.5(4)	159.5(8)	135.5(8)	140.0(6)	54.0(10)
CC+T	1	53.0(10)	52.0(10)	51.5(10)	51.0(10)	51.5(10)
	2	57.0(9)	54.0(10)	53.0(10)	52.0(10)	52.5(10)
	3	59.0(5)	56.0(9)	56.0(10)	56.0(10)	51.0(10)
AC	1	65.5(10)	61.0(10)	48.0(10)	41.5(10)	33.0(10)
	2	85.0(10)	64.0(10)	56.0(10)	46.0(10)	35.0(10)
	3	105.0(10)	70.5(10)	53.0(10)	46.5(10)	35.5(10)
AC+T	1	55.5(10)	59.0(10)	58.0(10)	57.5(10)	57.5(10)
	2	59.0(10)	60.0(10)	57.5(10)	60.5(10)	58.0(10)
	3	58.5(10)	58.5(10)	58.5(10)	59.0(10)	59.5(10)
k10	1	48.0(10)	<b>40.5(10)</b>	<b>33.5(10)*</b>	31.5(10)	<b>23.0(10)</b>
	2	50.0(10)	<b>41.5(10)*</b>	<b>38.0(10)*</b>	<b>31.0(10)*</b>	<b>26.5(10)</b>
	3	53.5(10)	47.0(10)	<b>37.5(10)*</b>	<b>34.0(10)</b>	<b>29.5(10)*</b>
k10+T	1	51.0(10)	52.0(10)	52.0(10)	52.5(10)	53.0(10)
	2	52.0(10)	53.0(10)	52.0(10)	53.0(10)	53.5(10)
	3	51.0(10)	53.0(10)	52.0(10)	52.5(10)	56.0(10)
SEM	1	47.0(10)	47.0(10)	42.5(10)	32.5(10)	34.0(10)
	2	50.0(10)	45.5(10)	41.5(10)	39.0(10)	32.5(10)
	3	51.5(10)	49.0(10)	43.0(10)	44.0(10)	41.5(10)
SEM+T	1	<b>46.0(10)</b>	44.5(10)	41.0(10)	<b>30.0(10)</b>	26.5(10)†
	2	49.5(10)	48.5(10)	42.0(10)	38.5(10)	33.5(10)
	3	<b>48.0(10)</b>	<b>44.0(10)†</b>	41.5(10)	41.5(10)	39.0(10)
SEM.bic	1	47.0(10)	46.0(10)	42.0(10)	38.5(10)	31.5(10)
	2	49.5(10)	48.5(10)	43.0(10)	44.5(10)	37.0(10)
	3	50.0(10)	50.0(10)	44.0(10)	44.5(10)	38.5(10)
SEM.bic+T	1	47.0(10)	45.5(10)	42.0(10)	40.0(10)	31.5(10)
	2	48.5(10)	46.0(10)	47.0(10)	42.0(10)	38.5(10)
	3	48.5(10)	46.0(10)	46.5(10)	43.0(10)	40.5(10)
GES-EM	1	47.5(10)	44.5(10)	39.5(10)	38.0(10)	-
	2	<b>48.0(10)</b>	45.5(10)	41.0(9)	42.0(7)	-
	3	51.0(10)	47.5(10)	44.5(10)	44.0(8)	-
GES-EM+T	1	52.0(9)	53.0(9)	53.0(10)	50.5(10)	-
	2	52.0(9)	53.0(10)	55.0(10)	54.0(10)	-
	3	53.0(9)	51.5(10)	53.0(10)	54.0(9)	-
SEM+RT	1	47.5(10)	45.0(10)	44.0(10)	36.0(10)	34.0(10)
	2	51.5(10)	48.0(10)	43.5(10)	41.5(10)	35.0(10)
	3	51.5(10)	50.0(10)	43.0(10)	46.0(10)	40.5(10)

SHD Insurance mar SHD incompleti

SHD Alarm1 mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>41.5(10)</i>	<i>34.0(10)</i>	<i>20.5(10)</i>	<i>20.0(10)</i>	<i>14.0(10)</i>
MWST-EM	1	55.5(10)	56.0(10)	56.5(10)	56.5(10)	57.0(10)
	2	58.0(10)	55.5(10)	59.0(10)	59.5(10)	56.0(10)
	3	56.5(10)	56.0(10)	56.5(10)	55.0(10)	58.5(10)
CC	1	279.5(6)	262.0(10)	60.0(10)	72.5(10)	31.5(10)
	2	130.5(4)	199.0(2)	245.5(10)	209.5(6)	60.0(10)
	3	132.5(2)	135.0(3)	280.0(1)	180.5(6)	206.0(7)
CC+T	1	50.0(10)	48.0(10)	46.5(10)	47.5(10)	47.5(10)
	2	58.0(5)	55.0(7)	52.0(10)	50.5(10)	46.0(10)
	3	57.0(3)	53.5(4)	53.0(1)	54.0(7)	54.5(8)
AC	1	96.0(10)	66.0(10)	49.5(10)	44.0(10)	38.0(10)
	2	125.0(10)	74.5(10)	53.5(10)	49.5(10)	39.5(10)
	3	151.5(10)	74.5(10)	58.0(10)	48.0(10)	37.5(10)
AC+T	1	50.5(10)	54.5(10)	52.0(10)	53.0(10)	58.5(10)
	2	55.0(10)	55.5(10)	53.5(10)	55.5(10)	59.5(10)
	3	53.5(10)	57.0(10)	53.5(10)	56.0(10)	58.5(10)
k10	1	46.5(10)	35.0(10)	27.0(10)	28.0(10)	22.5(10)
	2	49.0(10)	42.0(10)	29.5(10)	30.0(10)	24.5(10)
	3	52.0(10)	42.0(10)	34.5(10)	32.5(10)	32.0(10)
k10+T	1	44.0(10)	43.0(10)	46.0(10)	46.0(10)	54.5(10)
	2	45.5(10)	44.0(10)	45.5(10)	48.0(10)	54.0(10)
	3	45.0(10)	45.0(10)	45.0(10)	46.5(10)	53.0(10)
SEM	1	34.0(10)	39.5(10)	33.5(10)	25.5(10)	25.5(10)
	2	44.5(10)	42.5(10)	37.0(10)	37.0(10)	27.0(10)
	3	42.0(10)	42.5(10)	40.0(10)	34.5(10)	32.5(10)
SEM+T	1	<b>30.0(10)*†</b>	<b>24.0(10)*†</b>	<b>17.0(10)*†</b>	<b>17.0(10)*†</b>	<b>6.5(10)*†</b>
	2	<b>32.5(10)*†</b>	<b>29.0(9)*†</b>	<b>27.5(10)†</b>	<b>26.5(10)†</b>	<b>13.5(10)†</b>
	3	<b>39.0(10)*†</b>	<b>32.0(9)*†</b>	<b>31.0(10)*†</b>	<b>23.5(10)*†</b>	<b>33.0(10)†</b>
SEM.bic	1	39.5(10)	44.5(10)	40.0(10)	35.0(10)	32.0(10)
	2	45.0(10)	46.0(10)	36.0(10)	38.0(10)	35.5(10)
	3	45.5(10)	45.5(10)	41.0(10)	36.5(10)	36.0(10)
SEM.bic+T	1	41.0(10)	36.5(10)	29.5(10)	31.0(10)	14.5(10)
	2	44.5(10)	39.0(10)	32.5(10)	30.5(10)	17.5(10)
	3	44.0(10)	40.5(10)	34.5(10)	30.0(10)	<b>29.5(10)</b>
GES-EM	1	38.0(10)	42.0(10)	33.0(9)	33.0(3)	-
	2	46.5(10)	46.0(9)	33.0(9)	37.0(3)	-
	3	44.5(10)	44.5(10)	44.0(9)	29.0(1)	-
GES-EM+T	1	54.0(10)	53.0(10)	53.0(10)	55.0(9)	-
	2	53.0(10)	51.0(10)	54.0(10)	51.5(10)	-
	3	53.0(10)	51.5(10)	50.5(10)	53.0(9)	-
SEM+RT	1	<i>38.5(10)</i>	<i>42.5(10)</i>	<i>37.5(10)</i>	<i>33.5(10)</i>	<i>33.5(10)</i>
	2	<i>44.5(10)</i>	<i>49.0(9)</i>	<i>37.5(10)</i>	<i>38.5(10)</i>	<i>43.5(10)</i>
	3	<i>48.0(10)</i>	<i>48.0(10)</i>	<i>38.0(10)</i>	<i>38.5(10)</i>	<i>39.5(10)</i>

SHD Alarm1 mar SHD incompleti

SHD HailFinder mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	95.5(10)	100.5(10)	105.0(10)	103.0(10)	101.5(10)
MWST-EM	1	108.0(10)	110.0(10)	109.0(10)	106.0(10)	104.5(10)
	2	109.5(10)	110.0(10)	108.0(10)	107.0(10)	102.5(10)
	3	111.0(10)	108.0(10)	106.5(10)	105.5(10)	101.5(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	91.5(10)	99.0(10)	106.0(10)	108.0(10)*	110.5(10)
	2	93.0(10)	97.5(10)	106.0(10)	107.0(10)	110.0(10)
	3	92.0(10)	98.0(10)	106.0(10)	109.0(10)	112.5(10)
k10+T	1	<b>76.0(10)</b>	<b>79.5(10)</b>	<b>81.0(10)</b>	<b>81.0(10)</b>	<b>86.5(10)</b>
	2	<b>76.5(10)</b>	<b>77.5(10)</b>	<b>81.0(10)</b>	<b>82.0(10)</b>	<b>87.0(10)</b>
	3	<b>76.0(10)</b>	<b>79.0(10)</b>	<b>79.0(10)</b>	<b>83.0(10)</b>	<b>88.0(10)</b>
SEM	1	84.0(10)	89.0(10)	101.5(10)	111.0(10)	109.0(10)
	2	82.0(10)	88.0(10)	100.0(10)	111.0(10)	108.0(5)
	3	80.0(10)	85.5(10)	97.5(10)	107.5(10)	111.5(8)
SEM+T	1	86.0(10)*	90.0(10)*	102.0(10)*	111.0(10)	114.0(10)
	2	84.0(10)*	89.0(10)*	102.5(10)*	109.0(9)	114.0(9)
	3	82.0(10)*	87.5(10)*	98.5(10)*	107.5(10)	111.0(7)
SEM.bic	1	84.0(10)	90.5(10)	104.5(10)	110.5(10)	111.0(8)
	2	83.0(10)	89.0(10)	104.5(10)	109.0(10)	108.5(6)
	3	81.0(10)	88.0(10)	101.0(10)	107.0(10)	111.0(8)
SEM.bic+T	1	85.5(10)	91.0(10)	105.5(10)	108.5(10)	115.0(9)
	2	84.0(10)	91.0(10)	104.5(10)	109.5(10)	115.0(8)
	3	82.0(10)	90.0(10)	103.0(10)	108.5(10)	113.0(6)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	85.0(10)	90.5(10)	102.5(10)	111.5(10)	111.5(10)
	2	83.0(10)	88.5(10)	100.5(10)	109.5(10)	113.0(7)
	3	80.0(10)†	86.5(10)†	99.5(10)	107.5(10)	114.5(8)

SHD HailFinder mar SHD incompleti

SHD Hepar2 mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	171.5(10)	151.0(10)	121.0(10)	104.0(10)	70.5(10)
MWST-EM	1	175.5(10)	173.0(10)	171.5(10)	175.0(10)	173.0(10)
	2	174.5(10)	176.0(10)	173.0(10)	172.5(10)	171.0(10)
	3	176.0(10)	173.0(10)	176.5(10)	175.5(10)	176.0(6)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	175.5(10)	159.5(10)	128.5(10)	113.0(10)	80.0(10)
	2	179.0(10)	159.5(10)	133.5(10)	113.0(10)	91.0(10)
	3	186.5(10)	164.5(10)	131.5(10)	124.5(10)	98.5(10)
k10+T	1	125.0(10)	123.0(10)	124.0(10)	126.0(10)	123.5(10)
	2	125.5(10)	124.0(10)	125.0(10)	123.5(10)	122.5(10)
	3	126.0(10)	125.0(10)	127.0(10)	125.0(10)	125.5(6)
SEM	1	120.0(10)	116.5(10)	<b>104.0(10)</b>	<b>92.5(10)</b>	<b>71.5(10)</b>
	2	<b>120.0(10)</b>	118.5(10)	109.0(10)	96.5(10)	<b>73.0(10)</b>
	3	121.5(10)	118.5(10)	109.0(10)	<b>96.5(10)</b>	<b>77.0(10)</b>
SEM+T	1	<b>119.5(10)*</b>	<b>116.0(10)*</b>	104.5(10)*	94.5(10)*	78.0(10)
	2	120.0(10)*	<b>117.0(10)*</b>	<b>106.5(10)*</b>	97.0(10)*	76.0(10)*
	3	121.0(10)*	<b>117.0(10)*†</b>	108.5(10)*	98.0(10)*	80.5(6)*
SEM.bic	1	119.5(10)	117.0(10)	105.0(10)	95.0(10)	73.5(10)
	2	120.5(10)	117.5(10)	108.0(10)	<b>95.5(10)</b>	75.0(10)
	3	121.0(10)	118.5(10)	<b>106.5(10)</b>	97.5(10)	78.5(10)
SEM.bic+T	1	119.5(10)	116.5(10)	106.5(10)	96.0(10)	83.0(10)
	2	121.0(10)	118.0(10)	108.0(10)	99.0(10)	81.0(10)
	3	<b>120.0(10)</b>	117.5(10)	109.5(10)	101.0(10)	84.5(6)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	120.0(10)	117.5(10)	105.5(10)	93.0(10)	72.0(10)
	2	120.5(10)	117.0(10)	105.0(10)	96.0(10)	73.5(10)
	3	121.0(10)	119.0(10)	107.5(10)	98.5(10)	78.5(10)

SHD Hepar2 mar SHD incompleti

SHD Asia nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	6.5(10)	6.5(10)	6.0(10)	6.0(10)	5.0(10)
MWST-EM	1	10.0(10)	11.0(10)	10.0(10)	10.5(10)	10.5(10)
	2	10.5(10)	10.5(10)	10.0(10)	10.0(10)	10.0(10)
	3	11.0(10)	10.0(10)	10.0(10)	10.5(10)	10.0(10)
CC	1	8.0(10)	7.5(10)	6.0(10)	6.5(10)	5.0(10)
	2	7.5(10)	7.5(10)	6.0(10)	6.0(10)	6.0(10)
	3	10.0(10)	8.5(10)	7.0(10)	6.0(10)	6.0(10)
CC+T	1	7.0(10)	8.0(10)	7.0(10)	9.0(10)	9.0(10)
	2	8.0(10)	7.0(10)	7.0(10)	8.0(10)	8.5(10)
	3	8.0(10)	7.0(10)	8.0(10)	7.5(10)	9.0(10)
AC	1	7.0(10)	9.0(10)	7.0(10)	8.0(10)	5.0(10)
	2	7.5(10)	9.0(10)	7.0(10)	8.0(10)	5.5(10)
	3	8.0(10)	8.0(10)	8.5(10)	7.5(10)	6.0(10)
AC+T	1	9.0(10)	8.5(10)	10.0(10)	10.0(10)	9.0(10)
	2	9.5(10)	9.0(10)	10.0(10)	10.0(10)	9.0(10)
	3	10.0(10)	9.0(10)	9.0(10)	10.5(10)	9.0(10)
k10	1	7.0(10)	9.0(10)	6.0(10)	7.0(10)	5.5(10)
	2	6.5(10)	8.0(10)	6.0(10)	7.0(10)	6.0(10)
	3	7.5(10)	8.0(10)	7.0(10)	6.5(10)	6.0(10)
k10+T	1	7.0(10)	8.0(10)	8.0(10)	9.0(10)	9.5(10)
	2	7.5(10)	8.0(10)	7.5(10)	9.0(10)	9.5(10)
	3	8.0(10)	7.5(10)	8.5(10)	9.0(10)	9.0(10)
SEM	1	7.0(10)	5.5(10)	4.0(10)	3.5(10)	4.0(10)
	2	6.0(10)	6.5(10)	4.0(10)	4.0(10)	3.5(10)
	3	<b>6.5(10)</b>	7.5(10)	5.0(10)	4.0(10)	<b>2.0(10)</b>
SEM+T	1	6.0(10)	<b>5.0(10)*</b>	3.5(10)*	2.5(10)*	4.0(10)*
	2	<b>5.0(10)*</b>	<b>5.0(10)*</b>	4.0(10)	4.0(10)*	<b>2.0(10)*</b>
	3	6.5(10)	<b>5.5(10)*</b>	5.0(10)	4.0(10)*	3.0(10)*
SEM.bic	1	6.5(10)	6.0(10)	3.0(10)	2.5(10)	3.5(10)
	2	7.0(10)	6.0(10)	4.0(10)	4.0(10)	3.5(10)
	3	7.5(10)	7.5(10)	<b>4.5(10)</b>	3.5(10)	2.5(10)
SEM.bic+T	1	<b>5.5(10)</b>	5.0(10)	<b>2.5(10)</b>	<b>1.0(10)</b>	<b>1.5(10)</b>
	2	6.0(10)	5.5(10)	<b>3.0(10)</b>	3.0(10)	2.0(10)
	3	6.5(10)	6.0(10)	4.5(10)	<b>3.0(10)</b>	4.0(10)
GES-EM	1	6.0(10)	7.0(9)	3.0(9)	2.0(10)	2.0(10)
	2	6.5(10)	7.0(10)	3.5(10)	<b>2.0(10)</b>	2.0(10)
	3	8.0(10)	8.0(10)	5.0(10)	3.5(10)	4.0(10)
GES-EM+T	1	6.5(10)	8.0(10)	7.5(10)	8.0(10)	11.5(10)
	2	7.0(7)	8.0(10)	7.0(10)	7.5(10)	9.0(10)
	3	8.0(10)	6.5(10)	7.5(10)	7.0(10)	9.0(10)
SEM+RT	1	5.0(10)	5.0(10)	4.5(10)	3.5(10)	4.0(10)
	2	7.5(10)	5.0(10)	4.0(10)	3.0(10)	2.0(10)
	3	7.5(10)	6.5(10)	5.5(10)	4.0(10)	4.0(10)

SHD Asia nmar SHD incompleti

SHD Child nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	20.0(10)	13.0(10)	9.0(10)	6.0(10)	3.5(10)
MWST-EM	1	43.0(10)	43.0(10)	43.0(10)	43.0(10)	42.0(10)
	2	43.0(10)	43.0(10)	43.0(10)	41.5(10)	42.0(10)
	3	42.0(10)	42.0(10)	43.0(10)	42.5(10)	42.0(10)
CC	1	39.5(10)	38.5(10)	17.5(10)	12.0(10)	8.0(10)
	2	66.0(10)	48.0(10)	34.0(10)	21.0(10)	10.5(10)
	3	93.0(9)	62.5(10)	58.0(10)	31.0(10)	13.5(10)
CC+T	1	28.5(10)	28.0(10)	27.0(10)	29.0(10)	30.0(10)
	2	32.0(10)	30.0(10)	27.5(10)	26.5(10)	29.5(10)
	3	33.0(9)	31.0(10)	31.0(10)	28.0(10)	29.0(10)
AC	1	37.5(10)	29.0(10)	18.5(10)	17.0(10)	17.0(10)
	2	41.5(10)	29.5(10)	23.0(10)	20.5(10)	17.0(10)
	3	43.0(10)	29.5(10)	25.5(10)	23.0(10)	19.0(10)
AC+T	1	37.0(10)	37.0(10)	38.5(10)	39.0(10)	39.0(10)
	2	38.0(10)	36.5(10)	38.0(10)	39.5(10)	39.5(10)
	3	38.5(10)	38.5(10)	39.0(10)	40.0(10)	39.5(10)
k10	1	20.0(10)	<b>14.0(10)*</b>	<b>10.0(10)*</b>	<b>9.0(10)</b>	5.0(10)
	2	<b>19.0(10)*</b>	17.0(10)	<b>10.5(10)*</b>	<b>10.0(10)*</b>	<b>4.5(10)*</b>
	3	21.0(10)	<b>15.0(10)*</b>	<b>15.5(10)*</b>	<b>12.5(10)*</b>	<b>5.5(10)</b>
k10+T	1	28.0(10)	28.5(10)	29.5(10)	32.5(10)	37.0(10)
	2	27.5(10)	28.5(10)	30.0(10)	32.0(10)	38.0(10)
	3	27.0(10)	28.0(10)	30.5(10)	32.5(10)	37.5(10)
SEM	1	20.5(10)	16.5(10)	14.0(10)	10.0(10)	<b>4.0(10)</b>
	2	22.0(10)	17.5(10)	17.5(10)	13.0(10)	5.5(10)
	3	21.5(10)	17.0(10)	18.5(10)	14.0(10)	7.0(10)
SEM+T	1	<b>19.5(10)</b>	17.5(10)	15.0(10)	11.0(10)	4.0(10)
	2	22.0(10)	19.0(10)	18.0(10)	11.5(10)	8.0(10)
	3	22.0(10)	20.5(10)	21.5(10)	14.5(10)	10.0(10)
SEM.bic	1	21.5(10)	15.0(10)	13.5(10)	10.0(10)	4.5(10)
	2	22.0(10)	<b>16.5(10)</b>	16.5(10)	13.5(10)	4.5(10)
	3	<b>20.5(10)</b>	19.0(10)	19.5(10)	12.5(10)	6.0(10)
SEM.bic+T	1	20.0(10)	18.0(10)	14.0(10)	12.0(10)	5.0(10)
	2	21.0(10)	21.0(10)	18.0(10)	11.5(10)	11.0(10)
	3	21.5(10)	22.0(10)	20.0(10)	14.5(10)	8.5(10)
GES-EM	1	20.5(10)	16.0(10)	11.0(10)	9.0(10)	7.0(9)
	2	21.0(10)	17.0(10)	11.5(10)	11.0(10)	8.0(8)
	3	22.0(10)	16.5(10)	17.0(10)	13.0(10)	8.0(8)
GES-EM+T	1	27.0(10)	26.0(10)	29.0(9)	24.5(10)	25.0(9)
	2	26.0(10)	29.0(9)	29.5(10)	31.5(4)	26.0(4)
	3	-	-	-	-	28.0(7)
SEM+RT	1	21.0(10)	16.0(10)	14.0(10)	10.0(10)	4.5(10)
	2	19.5(10)	17.5(10)	19.0(10)	13.5(10)	3.5(10)†
	3	22.0(10)	19.5(10)	19.5(10)	17.0(10)	11.0(10)

SHD Child nmar SHD incompleti

SHD Insurance nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	42.5(10)	37.5(10)	31.5(10)	31.5(10)	21.0(10)
MWST-EM	1	64.0(10)	61.0(10)	62.0(10)	59.5(10)	57.0(10)
	2	64.5(10)	63.0(10)	62.0(10)	60.0(10)	59.5(10)
	3	63.5(10)	62.0(10)	62.0(10)	61.5(10)	58.5(10)
CC	1	110.0(10)	54.5(10)	45.5(10)	40.0(10)	31.0(10)
	2	176.0(7)	104.0(9)	75.5(10)	54.0(10)	36.5(10)
	3	139.0(4)	154.0(7)	128.0(8)	66.0(8)	59.5(10)
CC+T	1	56.0(10)	52.0(10)	52.0(10)	52.0(10)	53.0(10)
	2	57.0(10)	54.0(10)	52.5(10)	51.5(10)	52.5(10)
	3	59.0(9)	56.5(10)	55.5(10)	52.0(9)	53.0(10)
AC	1	70.0(10)	48.0(10)	43.5(10)	40.5(10)	31.5(10)
	2	92.0(10)	60.0(10)	48.0(10)	41.5(10)	34.5(10)
	3	105.0(10)	70.0(10)	51.5(10)	45.5(10)	37.5(10)
AC+T	1	57.0(10)	54.0(10)	57.5(10)	55.5(10)	58.5(10)
	2	58.0(10)	57.0(10)	58.0(10)	57.0(10)	61.5(10)
	3	59.0(10)	57.5(10)	58.5(10)	57.0(10)	59.0(10)
k10	1	50.0(10)	<b>39.5(10)</b>	<b>36.0(10)</b>	<b>32.0(10)*</b>	<b>23.0(10)</b>
	2	50.5(10)	47.0(10)*	<b>39.0(10)</b>	<b>35.5(10)</b>	<b>25.0(10)*</b>
	3	51.0(10)	49.5(10)	<b>40.5(10)</b>	<b>36.0(10)</b>	<b>30.0(10)</b>
k10+T	1	51.0(10)	52.5(10)	53.0(10)	54.0(10)	51.0(10)
	2	52.0(10)	54.0(10)	52.0(10)	52.5(10)	54.5(10)
	3	51.0(10)	53.0(10)	52.5(10)	53.0(10)	54.5(10)
SEM	1	<b>45.0(10)</b>	45.5(10)	40.5(10)	40.0(10)	29.0(10)
	2	49.0(10)	47.5(10)	42.0(10)	38.5(10)	33.0(10)
	3	<b>46.5(10)</b>	49.5(10)	44.0(10)	42.5(10)	40.0(10)
SEM+T	1	47.0(10)	45.0(10)	41.0(10)	38.0(10)	25.5(10)†
	2	<b>48.0(10)</b>	49.5(10)	42.0(10)	37.0(10)	33.0(8)
	3	46.5(10)*	<b>48.0(10)</b>	43.5(10)	42.5(10)	36.0(10)†
SEM.bic	1	48.5(10)	46.0(10)	43.0(10)	40.0(10)	34.0(10)
	2	50.0(10)	49.0(10)	44.5(10)	42.5(10)	39.0(10)
	3	49.5(10)	49.5(10)	45.0(10)	44.5(10)	40.5(10)
SEM.bic+T	1	49.0(10)	47.0(10)	45.5(10)	40.0(10)	33.0(10)
	2	49.0(10)	48.5(10)	48.0(10)	43.5(10)	37.0(10)
	3	48.5(10)	49.0(10)	47.5(10)	47.0(10)	34.5(10)
GES-EM	1	47.0(10)	44.5(10)	42.5(10)	43.0(7)	-
	2	49.5(10)	<b>46.0(9)</b>	45.0(9)	46.0(8)	-
	3	49.5(10)	48.5(10)	44.0(10)	45.0(10)	-
GES-EM+T	1	53.5(10)	52.0(9)	54.5(10)	54.5(10)	-
	2	50.5(10)	52.0(9)	54.5(10)	53.5(10)	-
	3	53.0(9)	52.5(10)	53.0(9)	53.5(10)	-
SEM+RT	1	47.5(10)	45.0(10)	40.0(10)	41.0(10)	34.0(10)
	2	49.5(10)	51.5(10)	40.0(10)	39.0(10)	37.0(10)
	3	51.0(10)	49.0(10)	47.0(10)	41.0(10)	41.0(10)

SHD Insurance nmar SHD incompleti

SHD Alarm1 nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>41.5(10)</i>	<i>34.0(10)</i>	<i>20.5(10)</i>	<i>20.0(10)</i>	<i>14.0(10)</i>
MWST-EM	1	57.5(10)	55.5(10)	56.0(10)	55.5(10)	55.5(10)
	2	60.5(10)	56.5(10)	56.5(10)	55.5(10)	59.5(10)
	3	61.5(10)	55.5(10)	56.0(10)	55.0(10)	57.5(10)
CC	1	266.5(10)	160.0(9)	164.0(9)	60.0(10)	31.5(10)
	2	187.0(5)	275.5(4)	300.0(5)	251.5(6)	50.0(10)
	3	134.0(4)	107.0(2)	197.5(2)	134.0(5)	240.0(7)
CC+T	1	52.5(10)	49.0(10)	49.0(10)	46.5(10)	45.5(10)
	2	54.0(7)	55.5(8)	52.0(10)	53.5(10)	46.0(10)
	3	60.5(4)	56.5(2)	56.5(2)	56.0(5)	51.0(9)
AC	1	103.5(10)	64.0(10)	49.5(10)	44.5(10)	35.5(10)
	2	122.0(10)	67.5(10)	55.0(10)	46.5(10)	34.5(10)
	3	147.5(10)	74.0(10)	53.5(10)	46.5(10)	38.0(10)
AC+T	1	51.0(10)	49.5(10)	54.5(10)	55.0(10)	55.0(10)
	2	53.5(10)	53.5(10)	55.0(10)	55.5(10)	56.0(10)
	3	54.5(10)	54.5(10)	57.0(10)	56.0(10)	56.0(10)
k10	1	46.5(10)	35.0(10)	31.5(10)	23.0(10)	24.0(10)
	2	51.0(10)	38.0(10)	31.5(10)	30.0(10)	24.5(10)
	3	52.0(10)	43.5(10)	38.0(10)	33.5(10)	31.0(10)
k10+T	1	44.0(10)	44.0(10)	47.0(10)	47.0(10)	51.5(10)
	2	46.0(10)	46.0(10)	47.0(10)	48.0(10)	51.5(10)
	3	48.0(10)	44.5(10)	47.0(10)	47.5(10)	51.5(10)
SEM	1	37.0(10)	37.5(10)	33.0(10)	27.0(10)	25.0(10)
	2	40.5(10)	41.5(10)	39.0(10)	34.0(10)	28.0(10)
	3	<b>40.5(10)</b>	41.0(10)	42.5(10)	39.0(10)	32.5(10)
SEM+T	1	<b>29.5(10)*†</b>	<b>27.5(10)*†</b>	<b>21.5(10)†</b>	<b>13.5(10)*†</b>	<b>8.5(10)*†</b>
	2	<b>36.5(10)*†</b>	<b>28.0(10)*†</b>	<b>23.0(10)*†</b>	<b>19.5(10)*†</b>	<b>23.0(10)†</b>
	3	40.5(10)*†	<b>31.5(10)*†</b>	<b>30.0(10)†</b>	<b>27.5(10)*†</b>	<b>25.0(10)†</b>
SEM.bic	1	39.0(10)	37.0(10)	38.0(10)	30.5(10)	30.5(10)
	2	45.5(10)	43.0(10)	40.0(10)	35.0(10)	33.5(10)
	3	43.5(10)	45.0(10)	42.5(10)	40.5(10)	38.0(10)
SEM.bic+T	1	41.0(10)	39.0(10)	33.0(10)	27.0(10)	15.0(10)
	2	46.5(10)	39.0(10)	29.0(10)	28.0(10)	31.0(10)
	3	46.5(10)	42.5(10)	34.5(10)	33.5(10)	27.0(10)
GES-EM	1	40.5(8)	40.0(9)	38.0(10)	32.0(3)	-
	2	42.0(9)	40.5(10)	41.5(10)	-	-
	3	44.0(9)	46.0(9)	43.0(10)	40.5(10)	-
GES-EM+T	1	53.5(10)	52.5(10)	51.0(10)	53.0(9)	-
	2	47.0(10)	54.0(10)	52.5(10)	55.0(10)	-
	3	50.5(10)	52.5(10)	53.0(10)	50.0(9)	-
SEM+RT	1	<i>40.5(10)</i>	<i>36.0(10)</i>	<i>37.0(10)</i>	<i>34.0(10)</i>	<i>32.5(10)</i>
	2	<i>45.5(10)</i>	<i>38.5(10)</i>	<i>43.0(10)</i>	<i>34.0(10)</i>	<i>38.5(10)</i>
	3	<i>43.5(10)</i>	<i>45.0(10)</i>	<i>39.5(10)</i>	<i>43.5(10)</i>	<i>40.0(10)</i>

SHD Alarm1 nmar SHD incompleti

SHD HailFinder nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>95.5(10)</i>	<i>100.5(10)</i>	<i>105.0(10)</i>	<i>103.0(10)</i>	<i>101.5(10)</i>
MWST-EM	1	110.0(10)	108.5(10)	108.0(10)	106.5(10)	101.5(10)
	2	110.5(10)	109.0(10)	110.0(10)	106.0(10)	105.0(10)
	3	110.0(10)	109.5(10)	108.0(10)	105.0(10)	103.5(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	90.0(10)	97.5(10)	104.0(10)	107.0(10)*	108.0(10)
	2	88.5(10)	96.5(10)	105.0(10)	108.5(10)	111.0(10)
	3	88.5(10)	94.5(10)	103.0(10)	108.0(10)	112.5(10)
k10+T	1	<b>77.5(10)</b>	<b>77.5(10)</b>	<b>80.5(10)</b>	<b>82.5(10)</b>	<b>88.0(10)</b>
	2	<b>75.0(10)</b>	<b>77.5(10)</b>	<b>81.5(10)</b>	<b>82.5(10)</b>	<b>87.0(10)</b>
	3	<b>76.5(10)</b>	<b>78.5(10)</b>	<b>81.5(10)</b>	<b>82.0(10)</b>	<b>88.5(10)</b>
SEM	1	84.5(10)	90.0(10)	102.5(10)	107.5(10)	109.0(7)
	2	81.5(10)	88.0(10)	100.5(10)	108.5(10)	111.0(5)
	3	80.0(10)	85.5(10)	96.0(10)	107.5(10)	114.0(7)
SEM+T	1	86.0(10)*	92.0(10)*	103.0(10)	110.5(10)	111.0(7)
	2	84.0(10)*	89.5(10)*	103.5(10)	111.0(10)	113.0(5)
	3	84.5(10)*	87.5(10)*	97.5(10)*	108.5(10)	115.0(6)
SEM.bic	1	82.5(10)	92.0(10)	104.0(10)	108.5(10)	111.0(8)
	2	81.0(10)	88.5(10)	104.5(10)	109.5(10)	113.0(5)
	3	79.5(10)	85.5(10)	99.0(10)	107.5(10)	112.5(8)
SEM.bic+T	1	83.5(10)	91.5(10)	105.0(10)	110.0(10)	111.5(8)
	2	83.5(10)	90.5(10)	104.5(10)	110.0(10)	110.0(5)
	3	82.5(10)	89.0(10)	101.0(10)	109.0(10)	112.0(7)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	<i>85.5(10)</i>	<i>91.5(10)</i>	<i>103.0(10)</i>	<i>110.0(10)</i>	<i>115.0(9)</i>
	2	<i>81.0(10)†</i>	<i>88.0(10)</i>	<i>103.0(10)</i>	<i>110.0(10)</i>	<i>115.0(8)</i>
	3	<i>82.0(10)†</i>	<i>87.0(10)</i>	<i>96.5(10)</i>	<i>107.5(10)</i>	<i>120.0(7)</i>

SHD HailFinder nmar SHD incompleti

SHD Hepar2 nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	171.5(10)	151.0(10)	121.0(10)	104.0(10)	70.5(10)
MWST-EM	1	172.5(10)	171.5(10)	174.5(10)	176.0(10)	174.0(10)
	2	174.0(10)	174.0(10)	172.5(10)	170.0(10)	174.5(10)
	3	174.5(10)	172.5(10)	175.0(10)	177.0(10)	173.0(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	174.0(10)	157.5(10)	124.0(10)	109.5(10)	81.0(10)
	2	182.0(10)	157.0(10)	130.5(10)	115.5(10)	82.0(9)
	3	186.5(10)	166.0(10)	137.0(10)	125.5(10)	87.0(9)
k10+T	1	124.5(10)	123.5(10)	124.5(10)	124.5(10)	126.5(10)
	2	126.0(10)	123.0(10)	123.0(10)	122.5(10)	127.0(10)
	3	126.5(10)	124.0(10)	125.0(10)	126.0(10)	123.5(10)
SEM	1	121.0(10)	116.5(10)	104.0(10)	<b>91.5(10)</b>	<b>68.5(10)</b>
	2	121.0(10)	118.0(10)	<b>106.0(10)</b>	96.5(10)	69.5(10)
	3	121.0(10)	119.5(10)	<b>109.0(10)</b>	<b>101.0(10)</b>	<b>76.0(10)</b>
SEM+T	1	120.5(10)*	118.0(10)*	105.5(10)*	97.0(10)*	76.5(10)
	2	121.0(10)*	118.0(10)*	107.5(10)*	<b>95.5(10)*</b>	76.5(10)*
	3	<b>120.5(10)*</b>	119.0(10)*	110.0(10)*	101.0(10)*	80.0(10)*
SEM.bic	1	120.0(10)	<b>116.0(10)</b>	103.5(10)	94.0(10)	74.0(10)
	2	<b>120.5(10)</b>	<b>117.5(10)</b>	108.5(10)	97.0(10)	<b>68.5(10)</b>
	3	121.5(10)	118.0(10)	109.5(10)	102.5(10)	78.5(10)
SEM.bic+T	1	<b>118.5(10)</b>	117.0(10)	<b>103.0(10)</b>	96.0(10)	80.5(10)
	2	120.5(10)	117.5(10)	107.5(10)	95.5(10)	81.5(10)
	3	121.0(10)	<b>117.5(10)</b>	111.0(10)	102.5(10)	81.5(10)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	119.5(10)	116.0(10)	103.0(10)†	93.5(10)	71.0(10)
	2	120.5(10)	117.0(10)	106.5(10)	97.5(10)	75.5(10)
	3	119.5(10)	119.0(10)	109.5(10)	100.0(10)	76.0(10)

SHD Hepar2 nmar SHD incompleti

**Precision**

Precision Asia mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	70.8(10)	62.5(10)	66.3(10)	83.3(10)	81.3(10)
	1	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
MWST-EM	2	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
	3	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
	1	50.0(10)	64.6(10)	69.0(10)	62.5(10)	76.7(10)
CC	2	20.0(10)	50.0(10)	66.7(10)	62.5(10)	83.3(10)
	3	31.0(10)	35.4(10)	50.0(10)	56.3(10)	62.5(10)
	1	<b>75.0(10)</b>	56.3(10)	64.6(10)	50.0(10)	40.8(10)
CC+T	2	68.8(10)	50.0(10)	56.3(10)	75.0(10)	50.0(10)
	3	<b>64.6(10)</b>	<b>75.0(10)</b>	75.0(10)	50.0(10)	43.8(10)
	1	35.4(10)	34.5(10)	45.8(10)	43.6(10)	70.0(10)
AC	2	43.8(10)	28.1(10)	39.6(10)	41.7(10)	50.0(10)
	3	35.4(10)	38.9(10)	41.7(10)	36.0(10)	50.0(10)
	1	40.0(10)	34.8(10)	41.7(10)	36.7(10)	31.7(10)
AC+T	2	40.0(10)	36.7(10)	36.7(10)	40.0(10)	40.0(10)
	3	40.0(10)	40.0(10)	40.0(10)	33.3(10)	40.0(10)
	1	50.0(10)	64.6(10)	62.5(10)	62.5(10)	78.9(10)
k10	2	40.8(10)	50.0(10)	64.6(10)	63.0(10)	83.8(10)
	3	37.5(10)	36.9(10)	51.9(10)	50.0(10)	66.7(10)
	1	75.0(10)	70.8(10)	50.0(10)	35.4(10)	38.8(10)
k10+T	2	<b>75.0(10)</b>	55.0(10)	45.8(10)	40.0(10)	37.5(10)
	3	62.5(10)	50.0(10)	50.0(10)	43.8(10)	36.7(10)
	1	58.3(10)	77.5(10)	88.9(10)	94.4(10)	<b>100.0(10)</b>
SEM	2	50.0(10)	81.7(10)	<b>88.9(10)</b>	88.9(10)	<b>100.0(10)</b>
	3	50.0(10)	43.8(10)	<b>83.3(10)</b>	<b>87.3(10)</b>	<b>100.0(10)</b>
	1	50.0(10)	70.8(10)	<b>100.0(10)*</b>	<b>100.0(10)*</b>	100.0(10)
SEM+T	2	50.0(10)	<b>82.9(10)*</b>	88.9(10)*	<b>94.4(10)*</b>	100.0(10)
	3	62.5(10)*	60.0(10)*	83.3(10)*	80.6(10)*	100.0(10)*
	1	50.0(10)	80.4(10)	94.4(10)	100.0(10)	100.0(10)
SEM.bic	2	50.0(10)	67.0(10)	87.3(10)	88.9(10)	100.0(10)
	3	50.0(10)	45.0(10)	64.6(10)	73.3(10)	100.0(10)
	1	62.5(10)	<b>84.5(10)</b>	100.0(10)	100.0(10)	100.0(10)
SEM.bic+T	2	50.0(10)	67.0(10)	80.2(10)	88.9(10)	100.0(10)
	3	56.3(10)	50.0(10)	68.3(10)	71.4(10)	100.0(10)
	1	54.2(10)	70.7(10)	88.9(10)	100.0(10)	100.0(10)
GES-EM	2	50.0(10)	60.0(10)	73.2(10)	94.4(10)	100.0(10)
	3	50.0(10)	40.8(10)	50.0(10)	63.3(10)	100.0(10)
	1	59.8(10)	56.3(10)	58.6(10)	60.4(10)	45.8(10)
GES-EM+T	2	55.0(10)	62.5(10)	58.6(10)	67.9(10)	73.9(10)
	3	56.3(10)	50.0(10)	50.0(10)	50.0(10)	66.7(10)
	1	58.3(10)	73.2(10)	88.9(10)	88.9(10)	91.7(10)
SEM+RT	2	50.0(10)	78.6(10)	94.4(10)	88.9(10)	100.0(10)
	3	33.3(10)	54.2(10)	84.5(10)	73.2(10)	100.0(10)

Precision Asia mcar Precisione in percentuale

Precision Child mcar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>63.5(10)</i>	<i>73.1(10)</i>	<i>80.0(10)</i>	<i>90.3(10)</i>	<i>100.0(10)</i>
MWST-EM	1	5.3(10)	7.9(10)	5.3(10)	5.3(10)	10.5(10)
	2	6.6(10)	5.3(10)	5.3(10)	5.3(10)	10.5(10)
	3	5.3(10)	5.3(10)	5.3(10)	6.6(10)	15.8(10)
CC	1	36.8(10)	63.1(10)	72.9(10)	80.0(10)	93.3(10)
	2	13.2(10)	14.4(10)	29.3(10)	45.1(10)	78.4(10)
	3	17.5(3)	14.0(4)	14.3(7)	12.7(10)	27.7(10)
CC+T	1	18.3(10)	29.2(10)	25.0(10)	20.0(10)	16.7(10)
	2	8.0(10)	11.7(10)	10.0(10)	22.5(10)	29.2(10)
	3	7.1(3)	4.1(4)	4.8(7)	8.7(10)	33.3(10)
AC	1	44.1(10)	65.0(10)	74.7(10)	84.4(10)	75.4(10)
	2	33.4(10)	40.4(10)	58.9(10)	60.6(10)	64.2(10)
	3	27.5(10)	37.8(10)	43.7(10)	48.8(10)	58.1(10)
AC+T	1	4.8(10)	7.7(10)	6.4(10)	6.1(10)	7.1(10)
	2	6.7(10)	6.1(10)	6.6(10)	6.3(10)	9.8(10)
	3	5.9(10)	6.0(10)	6.5(10)	8.4(10)	10.9(10)
k10	1	61.8(10)	70.3(10)	<b>78.1(10)</b>	<b>87.7(10)*</b>	<b>100.0(10)</b>
	2	55.7(10)	66.2(10)*	<b>75.3(10)*</b>	<b>81.7(10)*</b>	<b>100.0(10)*</b>
	3	53.2(10)	60.3(10)*	<b>70.8(10)*</b>	<b>75.0(10)*</b>	90.0(10)*
k10+T	1	22.5(10)	22.5(10)	18.3(10)	13.4(10)	15.0(10)
	2	29.2(10)	17.1(10)	15.5(10)	12.9(10)	17.5(10)
	3	18.3(10)	25.0(10)	13.4(10)	12.5(10)	18.8(10)
SEM	1	68.5(10)	<b>71.5(10)</b>	71.6(10)	83.6(10)	100.0(10)
	2	<b>62.5(10)</b>	64.4(10)	66.7(10)	69.4(10)	94.7(10)
	3	<b>75.0(10)</b>	63.7(10)	60.4(10)	64.7(10)	94.4(10)
SEM+T	1	<b>71.7(10)*</b>	66.7(10)	71.7(10)	82.2(10)	94.7(10)
	2	61.3(10)	57.2(10)	57.3(10)	73.6(10)	76.9(10)
	3	54.2(10)	50.0(10)	48.4(10)	57.4(10)	82.4(9)
SEM.bic	1	63.5(10)	68.9(10)	73.3(10)	84.5(10)	100.0(10)
	2	58.7(10)	<b>69.0(10)</b>	70.8(10)	73.1(10)	97.4(10)
	3	75.0(10)	<b>66.8(10)</b>	65.2(10)	68.5(10)	94.4(10)
SEM.bic+T	1	60.3(10)	61.2(10)	67.7(10)	81.9(10)	97.4(10)
	2	60.4(10)	57.2(10)	61.8(10)	74.3(10)	91.7(9)
	3	54.2(10)	51.9(10)	43.8(10)	59.7(10)	81.1(9)
GES-EM	1	62.5(10)	70.3(10)	73.5(10)	82.9(10)	91.7(10)
	2	53.6(10)	66.1(10)	69.4(10)	72.9(10)	83.3(9)
	3	66.7(10)	60.6(10)	63.9(10)	67.6(10)	<b>95.0(10)</b>
GES-EM+T	1	35.8(10)	46.2(9)	37.3(10)	50.0(10)	62.3(10)
	2	41.9(10)	31.5(10)	32.8(10)	42.9(6)	59.4(8)
	3	-	-	-	-	56.3(3)
SEM+RT	1	<i>69.6(10)</i>	<i>66.0(10)</i>	<i>69.4(10)</i>	<i>81.7(10)</i>	<i>85.4(10)</i>
	2	<i>68.5(10)</i>	<i>60.0(10)</i>	<i>61.4(10)</i>	<i>68.6(10)</i>	<i>91.7(10)†</i>
	3	<i>70.8(10)†</i>	<i>56.3(10)</i>	<i>53.1(10)</i>	<i>55.7(10)</i>	<i>81.1(10)</i>

Precision Child mcar Precisione in percentuale

Precision Insurance mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	60.6(10)	70.4(10)	76.5(10)	74.9(10)	89.0(10)
MWST-EM	1	24.0(10)	25.0(10)	27.9(10)	29.8(10)	34.6(10)
	2	23.1(10)	28.8(10)	28.8(10)	27.9(10)	32.7(10)
	3	27.9(10)	28.8(10)	29.8(10)	30.8(10)	32.7(10)
CC	1	38.4(10)	47.5(10)	70.2(10)	73.9(10)	82.8(10)
	2	11.5(5)	10.6(7)	11.8(10)	12.0(10)	52.2(10)
	3	-	17.3(3)	18.1(1)	14.6(1)	10.9(9)
CC+T	1	50.0(10)	45.8(10)	43.3(10)	44.1(10)	40.3(10)
	2	22.8(6)	27.0(10)	35.5(10)	30.4(10)	58.6(10)
	3	-	25.6(3)	28.2(1)	35.3(1)	33.0(10)
AC	1	40.6(10)	42.8(10)	56.8(10)	51.8(10)	51.3(10)
	2	20.2(10)	26.5(10)	45.7(10)	47.4(10)	53.6(10)
	3	16.5(10)	23.3(10)	35.0(10)	41.0(10)	56.7(10)
AC+T	1	33.4(10)	27.0(10)	29.7(10)	32.1(10)	30.6(10)
	2	25.3(10)	32.5(10)	29.5(10)	27.3(10)	29.9(10)
	3	31.6(10)	32.4(10)	30.1(10)	29.1(10)	30.4(10)
k10	1	<b>59.2(10)*</b>	<b>64.4(10)</b>	<b>72.6(10)*</b>	<b>82.3(10)</b>	<b>88.8(10)*</b>
	2	47.2(10)	<b>63.7(10)*</b>	<b>65.3(10)</b>	<b>71.3(10)*</b>	<b>81.5(10)</b>
	3	40.8(10)	<b>51.1(10)</b>	<b>57.2(10)*</b>	<b>61.9(10)</b>	<b>65.1(10)</b>
k10+T	1	53.6(10)	42.3(10)	40.5(10)	38.0(10)	38.6(10)
	2	43.7(10)	44.1(10)	43.2(10)	38.0(10)	40.2(10)
	3	48.3(10)	41.1(10)	40.8(10)	37.6(10)	38.8(10)
SEM	1	52.8(10)	62.6(10)	66.0(10)	71.6(10)	80.4(10)
	2	44.8(10)	51.2(10)	61.9(10)	61.2(10)	78.0(9)
	3	<b>50.0(10)</b>	46.7(10)	52.1(10)	59.6(10)	64.9(4)
SEM+T	1	51.9(10)	61.8(10)	57.0(10)	68.2(10)	81.6(10)
	2	45.5(10)	48.9(10)	57.5(10)	57.5(10)	72.7(7)
	3	48.3(10)*	46.3(10)	48.3(10)	52.2(10)	63.3(1)
SEM.bic	1	50.0(10)	53.5(10)	62.0(10)	59.6(10)	75.3(10)
	2	46.2(10)	50.0(10)	58.7(10)	58.7(10)	71.4(9)
	3	50.0(10)	47.2(10)	53.5(10)	58.2(10)	55.7(6)
SEM.bic+T	1	52.6(10)	53.5(10)	57.0(10)	66.7(10)	81.9(10)
	2	46.7(10)	50.0(10)	51.1(10)	55.6(10)	73.6(8)
	3	47.7(10)	46.1(10)	48.9(10)	50.0(10)	53.8(2)
GES-EM	1	49.9(10)	54.9(10)	62.5(9)	66.7(9)	-
	2	48.5(10)	55.0(10)	57.1(9)	61.4(9)	-
	3	40.0(9)	50.0(10)	53.2(10)	52.7(10)	-
GES-EM+T	1	37.5(10)	32.6(10)	41.9(10)	40.7(10)	-
	2	<b>50.0(10)</b>	45.8(10)	51.4(10)	47.3(10)	-
	3	50.0(10)	45.8(10)	34.1(10)	37.2(8)	-
SEM+RT	1	51.5(10)	55.3(10)	64.6(10)	72.0(10)	81.8(10)
	2	39.4(10)	50.0(10)	58.3(10)	55.8(10)	68.9(7)
	3	43.0(10)	47.5(10)	47.1(10)	43.9(9)	57.3(2)

Precision Insurance mcar Precisione in percentuale

Precision Alarm1 mcar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>45.6(10)</i>	<i>52.1(10)</i>	<i>66.4(10)</i>	<i>69.7(10)</i>	<i>72.1(10)</i>
MWST-EM	1	22.2(10)	22.2(10)	22.2(10)	22.2(10)	22.2(10)
	2	22.2(10)	22.2(10)	20.8(10)	22.2(10)	22.2(10)
	3	18.1(10)	20.8(10)	22.2(10)	22.2(10)	22.2(10)
CC	1	10.4(10)	15.2(10)	39.6(10)	54.5(10)	63.2(10)
	2	2.8(3)	3.7(4)	5.3(2)	5.6(5)	12.3(10)
	3	-	-	6.1(1)	7.8(1)	0.0(1)
CC+T	1	27.4(10)	29.1(10)	33.8(10)	38.1(10)	35.2(10)
	2	23.8(4)	20.3(6)	21.3(3)	21.6(10)	29.5(10)
	3	-	-	22.6(1)	19.6(1)	26.1(1)
AC	1	27.0(10)	31.9(10)	36.0(10)	45.2(10)	46.8(10)
	2	14.9(10)	25.3(10)	29.4(10)	39.0(10)	43.6(10)
	3	10.5(10)	20.8(10)	30.1(10)	35.5(10)	44.4(10)
AC+T	1	24.5(10)	27.6(10)	24.8(10)	25.5(10)	25.9(10)
	2	23.6(10)	23.8(10)	22.8(10)	25.7(10)	23.2(10)
	3	21.6(10)	23.5(10)	25.0(10)	23.8(10)	23.3(10)
k10	1	44.9(10)	46.5(10)	59.3(10)	61.5(10)	68.7(10)
	2	36.5(10)	46.2(10)	53.8(10)	56.3(10)	59.7(10)*
	3	28.3(10)	35.6(10)	45.1(10)	47.2(10)	46.3(10)
k10+T	1	36.9(10)	37.9(10)	37.5(10)	36.7(10)	26.2(10)
	2	33.4(10)	34.8(10)	38.0(10)	36.3(10)	25.6(10)
	3	31.9(10)	34.3(10)	37.5(10)	38.8(10)	29.9(10)
SEM	1	55.9(10)	62.4(10)	68.6(10)	70.0(10)	60.5(10)
	2	43.8(10)	50.0(10)	60.8(10)	58.5(9)	<b>69.1(8)</b>
	3	31.1(10)	<b>41.4(10)</b>	43.5(10)	59.6(8)	55.7(1)
SEM+T	1	<b>75.7(10)*†</b>	<b>75.7(10)*†</b>	<b>85.2(10)*†</b>	<b>85.1(10)*†</b>	<b>90.1(10)*†</b>
	2	<b>52.3(10)*†</b>	<b>59.8(10)*†</b>	<b>68.0(10)*†</b>	<b>70.2(10)*†</b>	44.5(8)
	3	<b>36.4(10)*†</b>	38.5(10)	<b>51.9(10)†</b>	<b>63.6(9)*†</b>	<b>65.5(5)</b>
SEM.bic	1	47.4(10)	45.5(10)	57.8(10)	56.8(10)	61.6(10)
	2	39.1(10)	40.9(10)	53.8(10)	50.1(10)	55.6(7)
	3	28.2(10)	35.4(10)	39.3(10)	58.8(10)	31.7(3)
SEM.bic+T	1	44.7(10)	49.1(10)	62.3(10)	64.9(10)	66.5(10)
	2	37.5(10)	43.6(10)	55.5(10)	60.2(10)	48.7(10)
	3	31.9(10)	39.3(10)	51.0(10)	51.8(9)	57.6(6)
GES-EM	1	45.5(10)	48.2(9)	59.5(10)	-	-
	2	40.2(10)	44.8(10)	54.8(10)	55.2(6)	-
	3	33.5(10)	37.2(10)	44.9(10)	55.9(10)	-
GES-EM+T	1	11.9(10)	22.1(10)	36.1(10)	35.0(7)	-
	2	25.5(10)	35.1(10)	-	48.1(3)	-
	3	20.0(10)	28.0(10)	36.4(9)	-	-
SEM+RT	1	<i>54.2(10)</i>	<i>53.3(10)</i>	<i>54.2(10)</i>	<i>66.4(10)</i>	<i>58.0(9)</i>
	2	<i>46.6(10)</i>	<i>41.3(10)</i>	<i>59.0(10)</i>	<i>50.0(9)</i>	<i>44.3(5)</i>
	3	<i>27.3(10)</i>	<i>35.7(10)</i>	<i>36.8(10)</i>	<i>44.9(7)</i>	-

Precision Alarm1 mcar Precisione in percentuale

Precision HailFinder mcar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	8.3(10)	7.7(10)	6.9(10)	8.5(10)	10.1(10)
MWST-EM	1	9.1(10)	8.6(10)	10.9(10)	11.4(10)	<b>20.0(10)</b>
	2	7.7(10)	8.2(10)	9.1(10)	11.8(10)	12.3(10)
	3	7.3(10)	10.0(10)	10.5(10)	12.3(10)	12.7(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	8.6(10)	8.6(10)	8.5(10)	7.6(10)	11.4(10)
	2	8.4(10)	8.8(10)	7.3(10)	8.7(10)	9.9(10)
	3	6.6(10)	7.6(10)	7.1(10)	6.5(10)	9.1(10)
k10+T	1	<b>14.7(10)</b>	<b>13.2(10)</b>	<b>17.3(10)</b>	<b>17.3(10)</b>	15.6(10)
	2	<b>13.6(10)</b>	<b>14.1(10)</b>	<b>16.0(10)</b>	<b>16.4(10)</b>	<b>15.4(10)</b>
	3	<b>13.8(10)</b>	<b>14.1(10)</b>	<b>15.3(10)</b>	<b>16.7(10)</b>	<b>18.3(10)</b>
SEM	1	7.8(10)	9.3(10)	10.1(10)	8.7(10)	11.9(8)
	2	6.9(10)	9.5(10)	8.6(10)	9.8(8)	7.6(1)
	3	0.0(10)	6.3(10)	7.8(10)	8.3(9)	-
SEM+T	1	9.2(10)	9.0(10)	9.6(10)	10.3(10)*†	14.6(8)*
	2	9.8(10)*†	10.0(10)	10.4(10)*	8.8(8)	10.2(2)
	3	12.9(10)*†	10.2(10)*†	10.5(10)*†	9.5(10)*	-
SEM.bic	1	8.7(10)	8.6(10)	9.1(10)	8.3(10)	15.2(5)
	2	7.7(10)	8.3(10)	8.8(10)	9.5(6)	-
	3	0.0(10)	7.3(10)	7.5(10)	7.3(7)	-
SEM.bic+T	1	10.6(10)	9.5(10)	8.8(10)	11.2(10)	14.2(6)
	2	9.3(10)	9.4(10)	9.1(10)	8.7(7)	-
	3	11.8(10)	9.3(10)	10.6(10)	9.8(9)	-
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	7.4(10)	8.1(10)	9.4(10)	8.3(10)	11.2(8)
	2	5.7(10)	9.8(10)	8.0(10)	9.6(8)	-
	3	0.0(10)	5.7(10)	7.3(9)	8.3(8)	-

Precision HailFinder mcar Precisione in percentuale

Precision Hepar2 mcar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	14.9(10)	22.1(10)	38.9(10)	53.2(10)	76.4(10)
MWST-EM	1	10.9(10)	12.3(10)	11.2(10)	12.3(10)	12.3(10)
	2	10.9(10)	12.3(10)	12.3(10)	12.3(10)	12.3(10)
	3	10.9(10)	12.3(10)	12.3(10)	12.3(10)	12.3(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	14.6(10)	20.5(10)	38.9(10)	50.6(10)	70.7(10)
	2	11.2(10)	20.1(10)	29.2(10)	41.4(10)	59.2(10)
	3	10.2(10)	15.0(10)	24.6(10)	34.5(10)	62.9(10)
k10+T	1	33.3(10)	39.4(10)	42.6(10)	40.9(10)	39.1(10)
	2	30.8(10)	40.3(10)	40.5(10)	42.6(10)	39.5(10)
	3	33.3(10)	36.1(10)	41.3(10)	40.9(10)	38.7(10)
SEM	1	<b>50.0(10)</b>	54.0(10)	60.0(10)	67.1(10)	80.2(10)
	2	50.0(10)	54.6(10)	<b>55.8(10)</b>	<b>64.1(10)</b>	82.2(10)
	3	50.0(10)	<b>57.7(10)</b>	53.2(10)	<b>57.3(10)</b>	76.8(10)
SEM+T	1	50.0(10)*	<b>54.5(10)*</b>	66.2(10)*	71.0(10)*	67.5(10)
	2	50.0(10)*	54.2(10)*	55.3(10)*	61.5(10)*	73.1(10)*
	3	<b>60.0(10)*</b>	50.0(10)*	50.0(10)*	56.4(10)*	70.1(10)
SEM.bic	1	50.0(10)	53.3(10)	<b>69.0(10)</b>	<b>71.4(10)</b>	<b>82.7(10)</b>
	2	50.0(10)	<b>55.1(10)</b>	55.2(10)	63.9(10)	<b>82.6(10)</b>
	3	50.0(10)	57.7(10)	<b>53.5(10)</b>	54.1(10)	<b>77.8(10)</b>
SEM.bic+T	1	50.0(10)	50.9(10)	66.8(10)	67.8(10)	65.5(10)
	2	<b>52.8(10)</b>	54.0(10)	55.1(10)	56.5(10)	75.9(10)
	3	59.2(10)	50.0(10)	47.3(10)	54.4(10)	66.0(10)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	50.0(10)	54.7(10)	60.2(10)	66.7(10)	79.0(10)†
	2	50.0(10)	54.8(10)	55.0(10)	63.2(10)	81.3(10)†
	3	50.0(10)	50.0(10)	50.0(10)	55.4(10)	72.7(10)

Precision Hepar2 mcar Precisione in percentuale

Precision Asia mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	70.8(10)	62.5(10)	66.3(10)	83.3(10)	81.3(10)
MWST-EM	1	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
	2	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
	3	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
CC	1	31.7(10)	62.5(10)	83.3(10)	77.4(10)	85.4(10)
	2	42.3(10)	66.7(10)	64.6(10)	66.3(10)	83.3(10)
	3	29.3(10)	31.7(10)	66.7(10)	68.3(10)	62.5(10)
CC+T	1	50.0(10)	64.6(10)	50.0(10)	50.0(10)	45.8(10)
	2	62.5(10)	62.5(10)	56.3(10)	56.3(10)	50.0(10)
	3	77.5(10)	61.3(10)	64.6(10)	75.0(10)	38.8(10)
AC	1	50.0(10)	40.8(10)	41.7(10)	50.0(10)	60.4(10)
	2	56.3(10)	40.0(10)	42.3(10)	47.7(10)	58.3(10)
	3	50.0(10)	33.3(10)	41.7(10)	41.7(10)	50.0(10)
AC+T	1	39.6(10)	45.8(10)	40.8(10)	38.2(10)	40.0(10)
	2	50.0(10)	37.5(10)	37.9(10)	36.7(10)	33.3(10)
	3	50.0(10)	41.7(10)	40.0(10)	36.7(10)	40.0(10)
k10	1	50.0(10)	60.4(10)	62.5(10)	83.3(10)	79.5(10)
	2	58.3(10)	61.3(10)	50.0(10)	71.4(10)	80.9(10)
	3	60.0(10)	50.0(10)	58.2(10)	52.3(10)	60.0(10)
k10+T	1	<b>70.8(10)</b>	62.5(10)	50.0(10)	45.0(10)	40.8(10)
	2	<b>75.0(10)</b>	50.0(10)	50.0(10)	45.0(10)	31.7(10)
	3	<b>79.2(10)</b>	61.3(10)	50.0(10)	43.8(10)	33.3(10)
SEM	1	69.0(10)	75.0(10)	88.9(10)	<b>100.0(10)</b>	<b>100.0(10)</b>
	2	53.6(10)	<b>80.0(10)</b>	77.4(10)	<b>94.4(10)</b>	87.5(10)
	3	43.8(10)	50.0(10)	66.7(10)	86.1(10)	82.2(10)
SEM+T	1	64.6(10)	<b>79.2(10)*</b>	<b>94.4(10)*</b>	94.4(10)*	100.0(10)
	2	66.7(10)	77.5(10)	<b>94.4(10)*</b>	88.9(10)*	<b>100.0(10)</b>
	3	52.8(10)	<b>62.5(10)*</b>	<b>87.3(10)*</b>	<b>88.9(10)*</b>	87.5(10)*
SEM.bic	1	55.0(10)	62.5(10)	88.9(10)	100.0(10)	80.0(10)
	2	50.0(10)	61.3(10)	84.5(10)	94.4(10)	100.0(10)
	3	43.8(10)	60.0(10)	70.7(10)	86.1(10)	90.0(10)
SEM.bic+T	1	50.0(10)	68.3(10)	88.9(10)	100.0(10)	100.0(10)
	2	62.5(10)	64.6(10)	88.9(10)	88.9(10)	100.0(10)
	3	56.3(10)	61.3(10)	85.7(10)	86.1(10)	<b>100.0(10)</b>
GES-EM	1	59.2(10)	59.2(10)	88.9(10)	92.9(10)	100.0(10)
	2	53.6(10)	60.0(10)	83.3(9)	94.4(10)	100.0(9)
	3	50.0(10)	55.0(10)	68.3(10)	55.0(10)	92.9(10)
GES-EM+T	1	62.5(9)	55.0(10)	45.8(10)	50.0(10)	43.8(10)
	2	61.3(10)	62.5(10)	52.8(10)	55.0(10)	46.2(9)
	3	50.0(9)	50.0(10)	50.0(10)	46.4(10)	44.3(10)
SEM+RT	1	69.0(10)	60.7(10)	100.0(10)	88.9(10)	100.0(10)
	2	69.0(10)	73.9(10)	78.6(10)	88.9(10)	100.0(10)
	3	50.0(10)	50.0(10)	76.2(10)	66.7(10)	69.7(10)

Precision Asia mar Precisione in percentuale

Precision Child mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>63.5(10)</i>	<i>73.1(10)</i>	<i>80.0(10)</i>	<i>90.3(10)</i>	<i>100.0(10)</i>
MWST-EM	1	5.3(10)	5.3(10)	5.3(10)	7.9(10)	10.5(10)
	2	6.6(10)	5.3(10)	5.3(10)	10.5(10)	13.2(10)
	3	5.3(10)	7.9(10)	6.6(10)	7.9(10)	10.5(10)
CC	1	27.4(10)	46.6(10)	73.6(10)	78.0(10)	88.4(10)
	2	12.2(10)	28.7(10)	40.9(10)	64.4(10)	77.0(10)
	3	11.7(9)	14.6(10)	18.4(10)	31.9(10)	65.6(10)
CC+T	1	15.0(10)	29.2(10)	25.0(10)	21.1(10)	17.9(10)
	2	7.2(10)	14.8(10)	20.0(10)	23.6(10)	23.2(10)
	3	7.7(9)	7.3(10)	12.5(10)	18.3(10)	18.3(10)
AC	1	33.3(10)	53.5(10)	53.5(10)	56.7(10)	66.5(10)
	2	28.4(10)	48.6(10)	51.8(10)	53.4(10)	66.7(10)
	3	28.5(10)	36.3(10)	47.4(10)	54.1(10)	63.1(10)
AC+T	1	10.0(10)	8.3(10)	6.7(10)	10.3(10)	15.1(10)
	2	7.3(10)	8.0(10)	6.6(10)	8.8(10)	11.3(10)
	3	8.8(10)	10.3(10)	10.0(10)	8.9(10)	12.7(10)
k10	1	61.1(10)	68.1(10)	<b>75.3(10)</b>	<b>86.0(10)*</b>	<b>100.0(10)*</b>
	2	56.5(10)	68.8(10)*	69.9(10)*	<b>78.6(10)*</b>	<b>97.4(10)</b>
	3	49.3(10)	66.6(10)	70.4(10)*	<b>78.2(10)*</b>	<b>91.0(10)*</b>
k10+T	1	27.5(10)	20.0(10)	14.3(10)	15.5(10)	16.0(10)
	2	27.5(10)	25.0(10)	13.4(10)	16.8(10)	20.9(10)
	3	21.8(10)	22.5(10)	16.7(10)	15.0(10)	14.5(10)
SEM	1	<b>64.0(10)</b>	71.8(10)	65.3(10)	80.6(10)	87.8(10)
	2	62.4(10)	70.9(10)	58.6(10)	73.0(10)	90.5(10)
	3	59.0(10)	68.3(10)	58.6(10)	67.8(10)	79.5(10)
SEM+T	1	64.0(10)	67.6(10)	67.7(10)	74.4(10)	94.7(10)
	2	<b>71.8(10)</b>	60.6(10)	59.7(10)	72.1(10)	82.5(10)
	3	56.3(10)	64.2(10)	64.1(10)	62.6(10)	70.6(10)
SEM.bic	1	55.0(10)	66.7(10)	68.1(10)	81.2(10)	87.4(10)
	2	63.1(10)	61.6(10)	62.5(10)	77.8(10)	92.0(10)
	3	56.3(10)	61.3(10)	60.3(10)	73.4(10)	85.1(10)
SEM.bic+T	1	55.0(10)	57.5(10)	69.9(10)	71.5(10)	97.1(10)
	2	54.5(10)	59.7(10)	59.7(10)	70.1(10)	93.2(10)
	3	53.9(10)	55.1(10)	61.4(10)	64.8(10)	84.0(10)
GES-EM	1	50.0(10)	<b>72.6(10)</b>	72.9(10)	81.4(10)	91.3(8)
	2	64.0(10)	<b>71.0(10)</b>	<b>70.6(10)</b>	74.3(10)	88.6(10)
	3	<b>60.4(10)</b>	<b>70.4(10)</b>	<b>72.4(10)</b>	72.2(10)	82.9(9)
GES-EM+T	1	41.9(10)	37.5(8)	38.2(9)	38.1(10)	54.8(9)
	2	28.4(10)	33.3(10)	35.7(10)	39.3(7)	60.0(6)
	3	-	-	-	-	42.4(7)
SEM+RT	1	<i>59.8(10)</i>	<i>66.7(10)</i>	<i>68.1(10)</i>	<i>73.8(10)</i>	<i>95.9(10)</i>
	2	<i>60.0(10)</i>	<i>62.7(10)†</i>	<i>55.8(10)</i>	<i>73.4(10)</i>	<i>94.4(10)</i>
	3	<i>60.4(10)</i>	<i>66.6(10)</i>	<i>56.1(10)</i>	<i>63.4(10)</i>	<i>78.1(10)</i>

Precision Child mar Precisione in percentuale

Precision Insurance mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	60.6(10)	70.4(10)	76.5(10)	74.9(10)	89.0(10)
MWST-EM	1	26.9(10)	27.9(10)	26.9(10)	30.8(10)	32.7(10)
	2	25.0(10)	27.9(10)	26.9(10)	26.9(10)	31.7(10)
	3	25.0(10)	30.8(10)	28.8(10)	30.8(10)	31.7(10)
CC	1	19.5(10)	33.5(10)	47.6(10)	60.7(10)	70.7(10)
	2	10.5(4)	21.4(9)	25.2(10)	37.5(9)	64.8(10)
	3	10.7(4)	9.5(8)	10.6(8)	10.3(6)	43.7(10)
CC+T	1	29.3(10)	50.0(10)	44.4(10)	43.0(10)	45.3(10)
	2	21.4(9)	33.3(10)	34.5(10)	45.3(10)	50.0(10)
	3	29.0(5)	20.0(9)	28.6(10)	26.1(10)	52.3(10)
AC	1	33.1(10)	38.1(10)	46.6(10)	57.8(10)	66.1(10)
	2	21.8(10)	35.5(10)	40.2(10)	47.3(10)	66.7(10)
	3	19.6(10)	32.2(10)	44.2(10)	47.8(10)	61.7(10)
AC+T	1	34.3(10)	30.2(10)	32.7(10)	35.1(10)	33.3(10)
	2	28.3(10)	30.1(10)	29.8(10)	28.8(10)	36.1(10)
	3	27.4(10)	31.5(10)	31.1(10)	31.4(10)	32.8(10)
k10	1	55.8(10)	<b>62.2(10)</b>	<b>80.2(10)*</b>	75.5(10)	<b>85.4(10)</b>
	2	50.0(10)	<b>63.6(10)*</b>	<b>69.0(10)</b>	<b>79.3(10)*</b>	<b>78.0(10)</b>
	3	48.8(10)	<b>58.3(10)</b>	<b>71.4(10)*</b>	<b>72.5(10)</b>	<b>72.8(10)*</b>
k10+T	1	49.1(10)	50.0(10)	38.8(10)	37.5(10)	36.4(10)
	2	42.9(10)	37.4(10)	40.0(10)	36.0(10)	35.7(10)
	3	50.0(10)	39.4(10)	37.3(10)	38.0(10)	33.9(10)
SEM	1	56.7(10)	52.0(10)	63.6(10)	76.9(10)	67.4(10)
	2	45.7(10)	49.4(10)	61.8(10)	66.3(10)	70.6(10)
	3	49.3(10)	48.2(10)	58.2(10)	56.2(10)	64.5(10)
SEM+T	1	<b>57.7(10)</b>	58.7(10)	64.5(10)	<b>80.6(10)</b>	79.2(10)
	2	49.3(10)	52.3(10)	63.6(10)	64.4(10)	71.6(10)
	3	54.1(10)*	55.7(10)	63.4(10)	60.8(10)	64.9(10)
SEM.bic	1	53.0(10)	50.0(10)	59.8(10)	64.4(10)	75.7(10)
	2	50.0(10)	50.0(10)	56.4(10)	54.4(10)	63.8(10)
	3	53.6(10)	44.9(10)	55.6(10)	56.9(10)	67.3(10)
SEM.bic+T	1	52.5(10)	53.6(10)	61.3(10)	63.0(10)	78.7(10)
	2	<b>53.2(10)</b>	51.1(10)	48.3(10)	63.3(10)	62.2(10)
	3	<b>54.3(10)</b>	51.2(10)	51.1(10)	57.3(10)	62.9(10)
GES-EM	1	51.2(10)	59.2(10)	63.8(10)	66.7(10)	-
	2	51.2(10)	52.4(10)	59.1(9)	61.9(7)	-
	3	48.6(10)	47.5(10)	57.5(10)	57.6(8)	-
GES-EM+T	1	39.7(9)	25.0(9)	38.8(10)	44.3(10)	-
	2	50.0(9)	35.6(10)	29.7(10)	40.0(10)	-
	3	30.8(9)	39.4(10)	34.4(10)	36.4(9)	-
SEM+RT	1	54.6(10)	57.3(10)	59.9(10)	70.7(10)	70.1(10)
	2	45.3(10)	51.8(10)	58.7(10)	63.2(10)	64.0(10)
	3	50.2(10)	46.1(10)	64.4(10)	50.0(10)	62.5(10)

Precision Insurance mar Precisione in percentuale

Precision Alarm1 mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>45.6(10)</i>	<i>52.1(10)</i>	<i>66.4(10)</i>	<i>69.7(10)</i>	<i>72.1(10)</i>
MWST-EM	1	22.2(10)	22.2(10)	22.2(10)	22.2(10)	22.2(10)
	2	20.1(10)	22.2(10)	22.2(10)	22.2(10)	22.2(10)
	3	22.2(10)	20.8(10)	22.2(10)	22.2(10)	22.2(10)
CC	1	8.2(6)	8.4(10)	31.4(10)	22.4(10)	56.8(10)
	2	3.6(4)	5.2(2)	7.6(10)	10.1(6)	30.2(10)
	3	2.9(2)	2.8(3)	5.8(1)	5.0(6)	4.1(7)
CC+T	1	25.0(10)	29.6(10)	32.3(10)	32.8(10)	35.9(10)
	2	22.2(5)	22.9(7)	23.4(10)	25.4(10)	33.4(10)
	3	25.0(3)	23.9(4)	22.2(1)	23.7(7)	25.1(8)
AC	1	18.2(10)	26.2(10)	37.7(10)	42.2(10)	45.0(10)
	2	13.0(10)	21.4(10)	32.8(10)	38.8(10)	43.9(10)
	3	10.5(10)	24.6(10)	29.1(10)	36.8(10)	44.5(10)
AC+T	1	29.0(10)	28.6(10)	27.8(10)	29.2(10)	23.6(10)
	2	26.0(10)	27.5(10)	26.9(10)	26.5(10)	21.3(10)
	3	26.5(10)	24.5(10)	26.9(10)	25.5(10)	22.7(10)
k10	1	40.2(10)	56.0(10)	64.4(10)	57.7(10)	64.6(10)
	2	40.4(10)	45.1(10)	56.8(10)	55.1(10)	60.6(10)
	3	36.0(10)	41.6(10)	51.8(10)	50.7(10)	51.7(10)
k10+T	1	36.7(10)	39.7(10)	35.0(10)	34.6(10)	26.8(10)
	2	32.3(10)	38.9(10)	38.2(10)	36.1(10)	25.0(10)
	3	36.8(10)	36.8(10)	38.6(10)	33.4(10)	27.4(10)
SEM	1	56.0(10)	50.7(10)	57.4(10)	65.7(10)	64.8(10)
	2	43.7(10)	45.2(10)	47.5(10)	53.4(10)	58.8(10)
	3	48.4(10)	43.8(10)	48.4(10)	58.9(10)	<b>55.8(10)</b>
SEM+T	1	<b>66.7(10)*†</b>	<b>73.9(10)*†</b>	<b>78.1(10)*†</b>	<b>78.5(10)*†</b>	<b>91.5(10)*†</b>
	2	<b>60.7(10)*†</b>	<b>64.6(9)*†</b>	<b>66.0(10)†</b>	<b>65.6(10)†</b>	<b>81.0(10)*†</b>
	3	<b>55.0(10)*†</b>	<b>65.8(9)*†</b>	<b>63.8(10)*†</b>	<b>67.0(10)*†</b>	52.7(10)
SEM.bic	1	41.5(10)	40.8(10)	46.2(10)	53.6(10)	51.8(10)
	2	36.5(10)	37.7(10)	49.8(10)	48.1(10)	50.2(10)
	3	35.7(10)	36.1(10)	46.9(10)	51.9(10)	46.9(10)
SEM.bic+T	1	43.6(10)	49.2(10)	59.5(10)	56.6(10)	75.2(10)
	2	38.0(10)	45.3(10)	53.2(10)	58.3(10)	71.5(10)
	3	36.3(10)	45.7(10)	53.2(10)	58.2(10)	55.0(10)
GES-EM	1	44.4(10)	41.2(10)	54.2(9)	57.8(3)	-
	2	35.1(10)	38.9(9)	56.1(9)	52.2(3)	-
	3	36.5(10)	39.7(10)	40.4(9)	58.3(1)	-
GES-EM+T	1	8.1(10)	17.5(10)	26.6(10)	25.0(9)	-
	2	11.1(10)	23.4(10)	26.6(10)	24.7(10)	-
	3	10.6(10)	27.0(10)	26.0(10)	22.2(9)	-
SEM+RT	1	<i>50.6(10)</i>	<i>43.3(10)</i>	<i>52.3(10)</i>	<i>55.4(10)</i>	<i>55.2(10)</i>
	2	<i>44.9(10)</i>	<i>35.3(9)</i>	<i>48.3(10)</i>	<i>51.6(10)</i>	<i>44.9(10)</i>
	3	<i>38.8(10)</i>	<i>36.3(10)</i>	<i>52.1(10)</i>	<i>50.0(10)</i>	<i>46.0(10)</i>

Precision Alarm1 mar Precisione in percentuale

Precision HailFinder mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	8.3(10)	7.7(10)	6.9(10)	8.5(10)	10.1(10)
MWST-EM	1	10.0(10)	9.1(10)	10.0(10)	10.9(10)	12.7(10)
	2	9.1(10)	8.2(10)	9.5(10)	10.9(10)	13.2(10)
	3	8.6(10)	10.0(10)	10.9(10)	12.7(10)	13.6(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	8.9(10)	8.0(10)	7.1(10)	7.7(10)	9.5(10)
	2	8.3(10)	7.8(10)	7.8(10)	7.9(10)	9.5(10)
	3	6.5(10)	7.0(10)	6.7(10)	7.4(10)	10.2(10)
k10+T	1	<b>15.3(10)</b>	<b>14.0(10)</b>	<b>15.6(10)</b>	<b>15.5(10)</b>	<b>15.2(10)</b>
	2	<b>13.8(10)</b>	<b>14.1(10)</b>	<b>15.7(10)</b>	<b>15.8(10)</b>	<b>14.3(10)</b>
	3	<b>12.5(10)</b>	<b>13.6(10)</b>	<b>17.7(10)</b>	<b>15.6(10)</b>	<b>14.8(10)</b>
SEM	1	6.2(10)	7.7(10)	8.1(10)	8.2(10)	11.8(10)
	2	5.3(10)	8.2(10)	7.8(10)	7.5(10)	13.6(5)
	3	3.8(10)	7.0(10)	6.9(10)	7.3(10)	11.1(8)
SEM+T	1	8.7(10)†	9.2(10)*†	9.0(10)*	10.0(10)*†	12.3(10)*
	2	9.8(10)*†	9.6(10)*†	9.0(10)	9.5(9)*	12.3(9)*
	3	9.5(10)*†	9.8(10)*†	10.3(10)*†	10.1(10)*†	13.4(7)*†
SEM.bic	1	7.0(10)	7.4(10)	7.9(10)	9.3(10)	11.8(8)
	2	6.1(10)	7.6(10)	6.7(10)	8.0(10)	12.7(6)
	3	4.5(10)	6.6(10)	6.7(10)	7.1(10)	10.1(8)
SEM.bic+T	1	9.3(10)	8.9(10)	8.5(10)	9.9(10)	11.9(9)
	2	9.6(10)	9.2(10)	8.4(10)	9.5(10)	11.1(8)
	3	10.3(10)	9.4(10)	9.2(10)	9.3(10)	10.8(6)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	7.1(10)	6.8(10)	7.9(10)	8.4(10)	13.4(10)
	2	5.0(10)	7.8(10)	7.6(10)	7.5(10)	10.3(7)
	3	5.4(10)	6.6(10)	6.6(10)	7.7(10)	9.9(8)

Precision HailFinder mar Precisione in percentuale

Precision Hepar2 mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>14.9(10)</i>	<i>22.1(10)</i>	<i>38.9(10)</i>	<i>53.2(10)</i>	<i>76.4(10)</i>
MWST-EM	1	10.9(10)	11.2(10)	12.3(10)	12.3(10)	12.7(10)
	2	10.9(10)	11.6(10)	11.6(10)	12.3(10)	12.7(10)
	3	10.9(10)	12.0(10)	12.3(10)	12.3(10)	12.3(6)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	12.7(10)	17.7(10)	34.9(10)	42.0(10)	66.1(10)
	2	13.3(10)	19.8(10)	31.1(10)	43.3(10)	52.2(10)
	3	11.4(10)	16.2(10)	33.5(10)	36.3(10)	52.4(10)
k10+T	1	36.0(10)	39.6(10)	41.6(10)	39.3(10)	39.4(10)
	2	32.6(10)	40.9(10)	41.3(10)	41.5(10)	39.7(10)
	3	32.1(10)	35.4(10)	40.1(10)	41.1(10)	37.5(6)
SEM	1	<b>50.0(10)</b>	<b>53.1(10)</b>	<b>60.0(10)</b>	66.9(10)	74.8(10)
	2	<b>50.0(10)</b>	<b>51.6(10)</b>	55.1(10)	67.4(10)	<b>78.7(10)</b>
	3	<b>50.0(10)</b>	51.1(10)	<b>58.5(10)</b>	62.7(10)	70.2(10)
SEM+T	1	50.0(10)*	49.1(10)*	56.3(10)*	59.2(10)*	64.5(10)
	2	50.0(10)*	51.4(10)*	<b>58.0(10)*</b>	62.1(10)*	67.1(10)*
	3	48.3(10)*	<b>52.7(10)*</b>	57.6(10)*	58.8(10)*	66.8(6)*
SEM.bic	1	50.0(10)	50.0(10)	58.5(10)	<b>69.8(10)</b>	<b>80.5(10)</b>
	2	50.0(10)	50.0(10)	54.2(10)	<b>72.8(10)</b>	75.2(10)
	3	50.0(10)	50.0(10)	57.1(10)	<b>63.9(10)</b>	<b>70.8(10)</b>
SEM.bic+T	1	50.0(10)	50.0(10)	58.1(10)	61.1(10)	60.3(10)
	2	50.0(10)	50.0(10)	57.0(10)	59.8(10)	64.7(10)
	3	50.0(10)	50.0(10)	51.1(10)	56.5(10)	57.1(6)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	<i>50.0(10)</i>	<i>51.5(10)</i>	<i>54.6(10)</i>	<i>67.1(10)</i>	<i>76.7(10)†</i>
	2	<i>44.4(10)</i>	<i>53.2(10)</i>	<i>59.0(10)</i>	<i>65.5(10)</i>	<i>74.7(10)</i>
	3	<i>50.0(10)</i>	<i>51.1(10)</i>	<i>56.8(10)</i>	<i>56.6(10)</i>	<i>67.3(10)</i>

Precision Hepar2 mar Precisione in percentuale

Precision Asia nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	70.8(10)	62.5(10)	66.3(10)	83.3(10)	81.3(10)
MWST-EM	1	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
	2	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
	3	35.7(10)	35.7(10)	35.7(10)	35.7(10)	35.7(10)
CC	1	58.3(10)	64.6(10)	70.8(10)	62.5(10)	87.5(10)
	2	53.6(10)	50.0(10)	72.9(10)	67.0(10)	70.0(10)
	3	36.9(10)	50.0(10)	61.3(10)	70.8(10)	70.7(10)
CC+T	1	62.5(10)	64.6(10)	50.0(10)	50.0(10)	50.0(10)
	2	50.0(10)	<b>68.8(10)</b>	56.3(10)	50.0(10)	50.0(10)
	3	<b>75.0(10)</b>	56.3(10)	50.0(10)	50.0(10)	50.0(10)
AC	1	50.0(10)	37.9(10)	45.8(10)	45.8(10)	70.0(10)
	2	45.0(10)	38.7(10)	50.0(10)	50.0(10)	58.3(10)
	3	35.4(10)	33.8(10)	46.4(10)	50.0(10)	58.3(10)
AC+T	1	50.0(10)	40.0(10)	38.8(10)	33.3(10)	37.5(10)
	2	45.8(10)	40.8(10)	40.0(10)	37.5(10)	45.0(10)
	3	40.8(10)	40.8(10)	38.8(10)	31.0(10)	40.0(10)
k10	1	66.7(10)	45.8(10)	61.3(10)	62.5(10)	68.3(10)
	2	66.7(10)	50.0(10)	72.9(10)	62.5(10)	65.2(10)
	3	58.3(10)	40.0(10)	55.6(10)	68.3(10)	76.4(10)
k10+T	1	<b>75.0(10)</b>	56.3(10)	50.0(10)	43.8(10)	38.8(10)
	2	68.8(10)	50.0(10)	50.0(10)	38.8(10)	31.7(10)
	3	50.0(10)	50.0(10)	37.5(10)	37.5(10)	33.3(10)
SEM	1	56.3(10)	64.6(10)	83.3(10)	88.9(10)	81.7(10)
	2	63.3(10)	58.3(10)	<b>88.9(10)</b>	88.9(10)	79.2(10)
	3	50.0(10)	50.0(10)	75.7(10)	<b>85.7(10)</b>	<b>100.0(10)</b>
SEM+T	1	61.4(10)	<b>68.3(10)</b>	88.2(10)*	94.4(10)*	81.7(10)*
	2	<b>72.5(10)</b>	66.7(10)	88.9(10)*	88.9(10)*	<b>100.0(10)*</b>
	3	50.0(10)†	<b>64.6(10)*</b>	73.2(10)	80.4(10)*	94.4(10)*
SEM.bic	1	56.3(10)	61.3(10)	<b>88.9(10)</b>	94.4(10)	90.0(10)
	2	50.0(10)	62.5(10)	88.9(10)	85.7(10)	82.9(10)
	3	50.0(10)	50.0(10)	76.7(10)	85.7(10)	92.9(10)
SEM.bic+T	1	62.5(10)	65.0(10)	88.9(10)	<b>100.0(10)</b>	<b>100.0(10)</b>
	2	62.5(10)	63.3(10)	88.9(10)	94.4(10)	100.0(10)
	3	50.0(10)	59.8(10)	<b>77.4(10)</b>	85.7(10)	88.9(10)
GES-EM	1	54.2(10)	50.0(9)	88.9(9)	100.0(10)	100.0(10)
	2	50.0(10)	50.0(10)	88.9(10)	<b>100.0(10)</b>	100.0(10)
	3	50.0(10)	45.8(10)	70.8(10)	85.7(10)	85.7(10)
GES-EM+T	1	53.6(10)	50.0(10)	40.8(10)	45.0(10)	39.6(10)
	2	50.0(7)	45.8(10)	50.0(10)	56.9(10)	43.8(10)
	3	50.0(10)	61.3(10)	50.0(10)	53.6(10)	46.9(10)
SEM+RT	1	71.4(10)	68.3(10)	76.2(10)	88.9(10)	80.0(10)
	2	41.4(10)	69.0(10)	87.3(10)	88.9(10)	100.0(10)
	3	45.0(10)	50.0(10)	65.7(10)	78.6(10)	81.7(10)

Precision Asia nmar Precisione in percentuale

Precision Child nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>63.5(10)</i>	<i>73.1(10)</i>	<i>80.0(10)</i>	<i>90.3(10)</i>	<i>100.0(10)</i>
MWST-EM	1	5.3(10)	5.3(10)	5.3(10)	5.3(10)	10.5(10)
	2	5.3(10)	5.3(10)	5.3(10)	10.5(10)	10.5(10)
	3	5.3(10)	7.9(10)	5.3(10)	7.9(10)	10.5(10)
CC	1	26.4(10)	30.0(10)	67.1(10)	79.7(10)	82.9(10)
	2	16.3(10)	22.3(10)	29.5(10)	61.8(10)	79.6(10)
	3	11.3(9)	15.3(10)	16.2(10)	38.4(10)	70.4(10)
CC+T	1	6.3(10)	23.6(10)	31.7(10)	20.0(10)	23.6(10)
	2	9.4(10)	12.7(10)	25.0(10)	29.2(10)	20.7(10)
	3	5.3(9)	10.3(10)	9.5(10)	25.0(10)	22.5(10)
AC	1	29.9(10)	42.0(10)	58.2(10)	60.1(10)	58.8(10)
	2	26.8(10)	39.8(10)	51.8(10)	54.3(10)	59.3(10)
	3	25.3(10)	40.3(10)	50.6(10)	50.0(10)	53.3(10)
AC+T	1	7.7(10)	8.5(10)	7.7(10)	8.5(10)	10.9(10)
	2	8.3(10)	8.2(10)	9.8(10)	11.4(10)	12.7(10)
	3	8.3(10)	10.7(10)	7.1(10)	8.7(10)	12.0(10)
k10	1	61.1(10)	68.5(10)	<b>76.7(10)*</b>	<b>82.2(10)*</b>	<b>94.8(10)</b>
	2	62.0(10)	64.7(10)	<b>75.3(10)*</b>	78.1(10)	<b>94.9(10)*</b>
	3	54.0(10)	67.2(10)	68.1(10)*	<b>79.2(10)*</b>	<b>92.2(10)*</b>
k10+T	1	25.0(10)	23.6(10)	18.3(10)	12.5(10)	16.9(10)
	2	29.2(10)	22.5(10)	17.7(10)	20.0(10)	14.8(10)
	3	26.7(10)	25.0(10)	16.7(10)	18.2(10)	16.7(10)
SEM	1	69.4(10)	71.4(10)	68.1(10)	76.7(10)	90.1(10)
	2	65.2(10)	67.4(10)	59.7(10)	68.6(10)	87.0(10)
	3	<b>72.5(10)</b>	68.8(10)	62.1(10)	70.1(10)	86.5(10)
SEM+T	1	<b>71.3(10)</b>	64.5(10)	61.1(10)	76.0(10)	93.3(10)
	2	<b>65.5(10)</b>	57.6(10)	58.6(10)	76.1(10)	83.1(10)
	3	65.5(10)*	61.6(10)	52.1(10)	67.1(10)	73.6(10)
SEM.bic	1	60.6(10)	70.7(10)	69.4(10)	80.0(10)	93.4(10)
	2	58.3(10)	67.4(10)	61.3(10)	69.1(10)	90.1(10)
	3	65.8(10)	64.1(10)	57.3(10)	75.7(10)	88.2(10)
SEM.bic+T	1	58.5(10)	59.7(10)	63.8(10)	76.5(10)	93.4(10)
	2	59.7(10)	53.5(10)	61.1(10)	<b>78.3(10)</b>	76.0(10)
	3	63.1(10)	57.7(10)	54.4(10)	71.2(10)	81.0(10)
GES-EM	1	57.0(10)	<b>73.4(10)</b>	73.5(10)	80.0(10)	83.8(9)
	2	58.9(10)	<b>68.5(10)</b>	73.5(10)	72.9(10)	84.0(8)
	3	62.4(10)	<b>72.2(10)</b>	<b>69.4(10)</b>	73.5(10)	86.2(8)
GES-EM+T	1	36.6(10)	41.7(10)	33.3(9)	46.6(10)	54.1(9)
	2	41.3(10)	33.3(9)	31.0(10)	31.4(4)	51.7(4)
	3	-	-	-	-	42.9(7)
SEM+RT	1	<i>67.5(10)</i>	<i>68.0(10)</i>	<i>66.7(10)</i>	<i>76.7(10)</i>	<i>93.3(10)</i>
	2	<i>66.6(10)</i>	<i>66.1(10)</i>	<i>58.8(10)</i>	<i>68.0(10)</i>	<i>93.4(10)†</i>
	3	<i>58.1(10)</i>	<i>63.1(10)</i>	<i>58.6(10)</i>	<i>60.8(10)</i>	<i>79.7(10)</i>

Precision Child nmar Precisione in percentuale

Precision Insurance nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	60.6(10)	70.4(10)	76.5(10)	74.9(10)	89.0(10)
MWST-EM	1	24.0(10)	30.8(10)	27.9(10)	28.8(10)	35.6(10)
	2	23.1(10)	26.0(10)	28.8(10)	30.8(10)	32.7(10)
	3	23.1(10)	28.8(10)	28.8(10)	29.8(10)	32.7(10)
CC	1	11.0(10)	45.5(10)	58.3(10)	63.1(10)	79.8(10)
	2	10.0(7)	16.5(9)	24.2(10)	46.6(10)	73.1(10)
	3	14.1(4)	10.4(7)	12.3(8)	30.5(8)	32.8(10)
CC+T	1	23.1(10)	46.9(10)	39.6(10)	50.0(10)	43.3(10)
	2	21.4(10)	29.2(10)	34.2(10)	50.0(10)	47.9(10)
	3	22.2(9)	25.5(10)	31.0(10)	42.9(9)	46.4(10)
AC	1	32.2(10)	51.9(10)	59.3(10)	59.0(10)	65.7(10)
	2	20.5(10)	38.2(10)	49.0(10)	56.1(10)	60.9(10)
	3	19.5(10)	33.1(10)	46.3(10)	51.6(10)	63.1(10)
AC+T	1	32.2(10)	36.6(10)	33.4(10)	33.4(10)	35.8(10)
	2	28.6(10)	31.9(10)	34.3(10)	34.8(10)	32.9(10)
	3	29.6(10)	33.9(10)	33.3(10)	33.5(10)	34.2(10)
k10	1	53.0(10)	<b>63.5(10)</b>	63.5(10)	<b>78.0(10)</b>	<b>85.1(10)</b>
	2	46.4(10)	<b>53.2(10)*</b>	60.8(10)	65.4(10)	<b>77.0(10)</b>
	3	48.8(10)	<b>52.2(10)</b>	59.7(10)	<b>65.2(10)</b>	<b>73.7(10)</b>
k10+T	1	50.0(10)	41.4(10)	39.4(10)	34.6(10)	40.0(10)
	2	<b>54.2(10)</b>	38.7(10)	40.9(10)	38.2(10)	36.2(10)
	3	50.0(10)	40.9(10)	40.5(10)	37.5(10)	34.8(10)
SEM	1	<b>54.9(10)</b>	55.1(10)	59.2(10)	66.8(10)	70.1(10)
	2	48.9(10)	51.1(10)	<b>63.2(10)</b>	70.6(10)	72.2(10)
	3	<b>57.5(10)</b>	48.1(10)	<b>59.8(10)</b>	59.0(10)	60.9(10)
SEM+T	1	48.7(10)	57.9(10)	<b>64.6(10)</b>	70.9(10)†	83.4(10)†
	2	48.6(10)	46.1(10)	59.5(10)	<b>72.4(10)</b>	68.1(8)
	3	55.2(10)*	50.0(10)	55.3(10)	63.4(10)	68.6(10)†
SEM.bic	1	50.0(10)	52.9(10)	55.0(10)	67.1(10)	68.6(10)
	2	47.0(10)	48.1(10)	55.2(10)	59.9(10)	65.3(10)
	3	50.0(10)	47.3(10)	55.3(10)	55.2(10)	62.8(10)
SEM.bic+T	1	48.5(10)	51.1(10)	54.3(10)	66.2(10)	67.9(10)
	2	48.7(10)	48.8(10)	49.4(10)	61.6(10)	60.6(10)
	3	48.7(10)	47.3(10)	52.1(10)	53.6(10)	70.8(10)
GES-EM	1	49.7(10)	54.7(10)	58.1(10)	58.0(7)	-
	2	46.5(10)	52.9(9)	54.5(9)	56.9(8)	-
	3	49.3(10)	50.0(10)	57.4(10)	56.4(10)	-
GES-EM+T	1	38.8(10)	22.5(9)	18.8(10)	28.2(10)	-
	2	50.0(10)	30.0(9)	22.6(10)	36.3(10)	-
	3	43.9(9)	38.8(10)	47.8(9)	40.9(10)	-
SEM+RT	1	50.0(10)	58.0(10)	62.0(10)	64.3(10)	70.3(10)
	2	46.4(10)	47.6(10)	60.5(10)	63.9(10)	61.7(10)
	3	46.1(10)	47.5(10)	56.0(10)	59.8(10)	57.0(10)

Precision Insurance nmar Precisione in percentuale

Precision Alarm1 nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>45.6(10)</i>	<i>52.1(10)</i>	<i>66.4(10)</i>	<i>69.7(10)</i>	<i>72.1(10)</i>
MWST-EM	1	20.1(10)	22.2(10)	22.2(10)	22.2(10)	22.2(10)
	2	18.8(10)	22.2(10)	22.2(10)	22.2(10)	22.2(10)
	3	19.4(10)	22.2(10)	22.2(10)	22.2(10)	22.2(10)
CC	1	8.3(10)	12.5(9)	10.4(9)	31.4(10)	55.9(10)
	2	4.0(5)	5.8(4)	5.4(5)	6.3(6)	38.7(10)
	3	2.5(4)	0.9(2)	6.7(2)	4.2(5)	9.1(7)
CC+T	1	23.2(10)	29.4(10)	26.8(10)	33.3(10)	36.8(10)
	2	22.4(7)	21.3(8)	26.0(10)	22.7(10)	32.0(10)
	3	16.6(4)	21.4(2)	20.6(2)	22.2(5)	26.7(9)
AC	1	17.1(10)	29.9(10)	34.7(10)	40.6(10)	47.9(10)
	2	14.2(10)	26.8(10)	33.5(10)	39.1(10)	47.4(10)
	3	10.5(10)	24.1(10)	32.8(10)	40.9(10)	42.0(10)
AC+T	1	26.1(10)	29.2(10)	28.4(10)	26.0(10)	24.5(10)
	2	24.5(10)	25.5(10)	26.0(10)	24.6(10)	22.3(10)
	3	23.1(10)	25.0(10)	24.5(10)	25.7(10)	23.9(10)
k10	1	39.2(10)	55.1(10)	55.1(10)	66.2(10)	61.4(10)
	2	40.5(10)	50.0(10)	56.1(10)	56.8(10)	59.7(10)
	3	34.6(10)	41.7(10)	45.4(10)	51.9(10)	52.2(10)
k10+T	1	33.9(10)	36.1(10)	36.8(10)	34.6(10)	28.0(10)
	2	32.4(10)	32.5(10)	36.5(10)	34.2(10)	29.2(10)
	3	32.9(10)	35.7(10)	32.5(10)	33.4(10)	29.8(10)
SEM	1	54.0(10)	48.3(10)	53.6(10)	63.8(10)	67.0(10)
	2	48.2(10)	46.2(10)	49.1(10)	53.6(10)	63.4(10)
	3	45.4(10)	48.2(10)	46.1(10)	48.6(10)	56.5(10)
SEM+T	1	<b>65.3(10)*†</b>	<b>69.4(10)*†</b>	<b>74.2(10)*†</b>	<b>84.5(10)*†</b>	<b>83.2(10)*†</b>
	2	<b>55.4(10)*†</b>	<b>66.7(10)*†</b>	<b>72.7(10)*†</b>	<b>72.3(10)*†</b>	<b>66.1(10)†</b>
	3	<b>51.0(10)*†</b>	<b>66.7(10)*†</b>	<b>62.9(10)*†</b>	<b>62.5(10)*†</b>	<b>63.1(10)†</b>
SEM.bic	1	46.5(10)	49.1(10)	49.2(10)	57.2(10)	52.4(10)
	2	37.2(10)	43.2(10)	45.6(10)	53.1(10)	51.2(10)
	3	37.8(10)	38.2(10)	39.6(10)	46.8(10)	44.9(10)
SEM.bic+T	1	40.8(10)	49.0(10)	55.3(10)	59.8(10)	74.1(10)
	2	36.0(10)	45.1(10)	59.1(10)	60.2(10)	54.1(10)
	3	34.4(10)	42.4(10)	52.0(10)	54.2(10)	59.7(10)
GES-EM	1	42.3(8)	47.3(9)	52.2(10)	56.3(3)	-
	2	38.6(9)	41.5(10)	45.2(10)	-	-
	3	40.0(9)	40.0(9)	41.9(10)	47.9(10)	-
GES-EM+T	1	13.7(10)	20.3(10)	23.8(10)	25.0(9)	-
	2	26.4(10)	15.0(10)	22.5(10)	17.4(10)	-
	3	18.8(10)	22.5(10)	20.4(10)	22.7(9)	-
SEM+RT	1	<i>50.0(10)</i>	<i>51.8(10)</i>	<i>51.0(10)</i>	<i>54.7(10)</i>	<i>54.3(10)</i>
	2	<i>42.6(10)</i>	<i>49.5(10)</i>	<i>42.2(10)</i>	<i>52.7(10)</i>	<i>45.1(10)</i>
	3	<i>40.4(10)</i>	<i>42.0(10)</i>	<i>47.2(10)</i>	<i>42.6(10)</i>	<i>44.4(10)</i>

Precision Alarm1 nmar Precisione in percentuale

Precision HailFinder nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	8.3(10)	7.7(10)	6.9(10)	8.5(10)	10.1(10)
MWST-EM	1	9.1(10)	10.0(10)	9.5(10)	10.5(10)	<b>16.8(10)</b>
	2	7.3(10)	9.1(10)	9.5(10)	11.8(10)	11.8(10)
	3	8.2(10)	9.5(10)	10.0(10)	12.7(10)	12.7(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	8.8(10)	8.2(10)	7.7(10)	8.0(10)	9.0(10)
	2	8.0(10)	8.1(10)	7.7(10)	7.2(10)	9.7(10)
	3	7.5(10)	7.6(10)	7.3(10)	6.1(10)	7.8(10)
k10+T	1	<b>14.3(10)</b>	<b>15.1(10)</b>	<b>16.0(10)</b>	<b>15.7(10)</b>	15.3(10)
	2	<b>15.0(10)</b>	<b>14.9(10)</b>	<b>15.2(10)</b>	<b>15.2(10)</b>	<b>14.6(10)</b>
	3	<b>11.7(10)</b>	<b>13.2(10)</b>	<b>16.3(10)</b>	<b>16.5(10)</b>	<b>14.0(10)</b>
SEM	1	8.4(10)	7.1(10)	8.1(10)	6.6(10)	11.0(7)
	2	5.9(10)	6.1(10)	7.6(10)	8.4(10)	11.6(5)
	3	6.5(10)	7.2(10)	7.0(10)	7.0(10)	9.2(7)
SEM+T	1	9.2(10)	8.3(10)†	9.0(10)	9.5(10)*†	13.5(7)
	2	8.9(10)†	9.1(10)†	9.1(10)*	9.0(10)	13.1(5)
	3	9.3(10)*†	9.8(10)*†	10.1(10)*†	9.2(10)*†	8.7(6)*
SEM.bic	1	10.2(10)	6.6(10)	7.7(10)	7.7(10)	12.9(8)
	2	6.5(10)	6.5(10)	6.4(10)	7.2(10)	13.6(5)
	3	6.5(10)	6.5(10)	7.8(10)	6.5(10)	9.4(8)
SEM.bic+T	1	10.6(10)	8.7(10)	8.2(10)	8.8(10)	12.4(8)
	2	10.2(10)	9.1(10)	8.4(10)	9.1(10)	12.9(5)
	3	9.0(10)	9.6(10)	8.7(10)	8.8(10)	10.4(7)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	7.5(10)	7.0(10)	8.3(10)	8.1(10)	11.0(9)
	2	5.7(10)	7.4(10)	6.8(10)	8.3(10)	11.4(8)
	3	5.9(10)	6.9(10)	8.4(10)	7.5(10)	7.3(7)

Precision HailFinder nmar Precisione in percentuale

Precision Hepar2 nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>14.9(10)</i>	<i>22.1(10)</i>	<i>38.9(10)</i>	<i>53.2(10)</i>	<i>76.4(10)</i>
MWST-EM	1	12.3(10)	11.6(10)	12.3(10)	11.2(10)	12.3(10)
	2	11.6(10)	12.0(10)	12.3(10)	12.3(10)	12.3(10)
	3	11.2(10)	12.0(10)	12.3(10)	12.3(10)	12.3(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	12.4(10)	21.2(10)	35.5(10)	47.5(10)	58.6(10)
	2	11.9(10)	19.7(10)	34.0(10)	44.7(10)	60.4(9)
	3	10.9(10)	16.8(10)	28.3(10)	35.8(10)	60.0(9)
k10+T	1	36.9(10)	40.2(10)	43.0(10)	39.3(10)	38.2(10)
	2	32.1(10)	43.3(10)	43.7(10)	40.5(10)	40.8(10)
	3	31.2(10)	41.9(10)	43.8(10)	39.3(10)	37.8(10)
SEM	1	<b>50.0(10)</b>	50.9(10)	59.5(10)	68.4(10)	<b>80.9(10)</b>
	2	50.0(10)	50.0(10)	<b>62.4(10)</b>	62.9(10)	82.1(10)
	3	50.0(10)	50.0(10)	<b>59.5(10)</b>	62.0(10)	<b>78.4(10)</b>
SEM+T	1	50.0(10)*	50.9(10)*	51.6(10)*	61.5(10)*	73.0(10)
	2	51.9(10)*	51.8(10)*	59.1(10)*	67.3(10)*	71.8(10)*
	3	50.0(10)*	<b>53.1(10)*†</b>	54.7(10)*	63.1(10)*	68.8(10)*
SEM.bic	1	50.0(10)	<b>52.6(10)</b>	61.6(10)	<b>70.3(10)</b>	72.5(10)
	2	<b>52.8(10)</b>	<b>53.0(10)</b>	55.2(10)	64.0(10)	<b>83.7(10)</b>
	3	<b>54.2(10)</b>	52.1(10)	58.8(10)	<b>68.4(10)</b>	76.2(10)
SEM.bic+T	1	50.0(10)	52.5(10)	<b>64.6(10)</b>	61.6(10)	69.0(10)
	2	52.5(10)	51.8(10)	55.2(10)	<b>69.9(10)</b>	64.8(10)
	3	50.0(10)	51.8(10)	53.8(10)	67.5(10)	68.7(10)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	<i>50.0(10)</i>	<i>52.6(10)</i>	<i>66.4(10)†</i>	<i>62.5(10)</i>	<i>77.7(10)</i>
	2	<i>50.0(10)</i>	<i>50.0(10)</i>	<i>58.9(10)</i>	<i>64.1(10)</i>	<i>74.5(10)</i>
	3	<i>50.0(10)</i>	<i>51.0(10)</i>	<i>57.3(10)</i>	<i>69.1(10)</i>	<i>71.5(10)</i>

**Precision Hepar2 nmar** Precisione in percentuale

## Recall

Recall Asia mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	36.4(10)	45.5(10)	45.5(10)	45.5(10)	72.7(10)
	1	45.5(10)	45.5(10)	45.5(10)	45.5(10)	45.5(10)
MWST-EM	2	<b>45.5(10)</b>	45.5(10)	45.5(10)	45.5(10)	45.5(10)
	3	<b>45.5(10)</b>	<b>45.5(10)</b>	45.5(10)	45.5(10)	45.5(10)
	1	18.2(10)	36.4(10)	45.5(10)	45.5(10)	63.6(10)
CC	2	9.1(10)	18.2(10)	40.9(10)	45.5(10)	45.5(10)
	3	40.9(10)	27.3(10)	18.2(10)	27.3(10)	45.5(10)
	1	40.9(10)	40.9(10)	31.8(10)	27.3(10)	27.3(10)
CC+T	2	27.3(10)	31.8(10)	27.3(10)	27.3(10)	27.3(10)
	3	22.7(10)	27.3(10)	31.8(10)	27.3(10)	27.3(10)
	1	31.8(10)	31.8(10)	45.5(10)	45.5(10)	63.6(10)
AC	2	27.3(10)	31.8(10)	45.5(10)	45.5(10)	45.5(10)
	3	18.2(10)	27.3(10)	45.5(10)	45.5(10)	59.1(10)
	1	36.4(10)	36.4(10)	45.5(10)	36.4(10)	31.8(10)
AC+T	2	36.4(10)	36.4(10)	36.4(10)	36.4(10)	36.4(10)
	3	36.4(10)	36.4(10)	36.4(10)	36.4(10)	36.4(10)
	1	31.8(10)	45.5(10)	45.5(10)	45.5(10)	77.3(10)
k10	2	36.4(10)	36.4(10)	59.1(10)	54.5(10)	77.3(10)
	3	27.3(10)	31.8(10)	54.5(10)	50.0(10)	<b>81.8(10)</b>
	1	31.8(10)	27.3(10)	27.3(10)	27.3(10)	27.3(10)
k10+T	2	45.5(10)	27.3(10)	27.3(10)	27.3(10)	27.3(10)
	3	31.8(10)	27.3(10)	27.3(10)	27.3(10)	36.4(10)
	1	36.4(10)	<b>54.5(10)</b>	77.3(10)	77.3(10)	81.8(10)
SEM	2	22.7(10)	45.5(10)	<b>72.7(10)</b>	72.7(10)	<b>81.8(10)</b>
	3	18.2(10)	27.3(10)	50.0(10)	<b>72.7(10)</b>	81.8(10)
	1	31.8(10)	54.5(10)	77.3(10)*	<b>81.8(10)*</b>	81.8(10)
SEM+T	2	27.3(10)	45.5(10)	72.7(10)*	<b>77.3(10)*</b>	81.8(10)
	3	36.4(10)†	45.5(10)	<b>63.6(10)</b>	68.2(10)	81.8(10)
	1	31.8(10)	54.5(10)	77.3(10)	81.8(10)	81.8(10)
SEM.bic	2	22.7(10)	45.5(10)	72.7(10)	72.7(10)	81.8(10)
	3	18.2(10)	27.3(10)	54.5(10)	54.5(10)	81.8(10)
	1	45.5(10)	54.5(10)	<b>81.8(10)</b>	81.8(10)	81.8(10)
SEM.bic+T	2	27.3(10)	45.5(10)	63.6(10)	72.7(10)	81.8(10)
	3	31.8(10)	45.5(10)	63.6(10)	50.0(10)	81.8(10)
	1	<b>50.0(10)</b>	54.5(10)	72.7(10)	81.8(10)	81.8(10)
GES-EM	2	31.8(10)	<b>54.5(10)</b>	63.6(10)	77.3(10)	81.8(10)
	3	36.4(10)	40.9(10)	45.5(10)	63.6(10)	81.8(10)
	1	45.5(10)	40.9(10)	59.1(10)	45.5(10)	<b>100.0(10)</b>
GES-EM+T	2	45.5(10)	45.5(10)	54.5(10)	59.1(10)	72.7(10)
	3	40.9(10)	45.5(10)	45.5(10)	45.5(10)	72.7(10)
	1	36.4(10)	54.5(10)	77.3(10)	72.7(10)	81.8(10)
SEM+RT	2	22.7(10)	45.5(10)	77.3(10)	72.7(10)	81.8(10)
	3	18.2(10)	45.5(10)	54.5(10)	54.5(10)	81.8(10)

Recall Asia mcar Recall in percentuale

Recall Child mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	45.9(10)	66.2(10)	77.0(10)	89.2(10)	100.0(10)
MWST-EM	1	5.4(10)	8.1(10)	5.4(10)	5.4(10)	10.8(10)
	2	6.8(10)	5.4(10)	5.4(10)	5.4(10)	10.8(10)
	3	5.4(10)	5.4(10)	5.4(10)	6.8(10)	16.2(10)
CC	1	28.4(10)	44.6(10)	52.7(10)	75.7(10)	91.9(10)
	2	27.0(10)	23.0(10)	23.0(10)	29.7(10)	64.9(10)
	3	<b>48.6(3)</b>	39.2(4)	45.9(7)	39.2(10)	24.3(10)
CC+T	1	2.7(10)	5.4(10)	5.4(10)	5.4(10)	10.8(10)
	2	2.7(10)	2.7(10)	1.4(10)	4.1(10)	5.4(10)
	3	2.7(3)	2.7(4)	2.7(7)	2.7(10)	5.4(10)
AC	1	<b>48.6(10)</b>	55.4(10)	60.8(10)	75.7(10)	62.2(10)
	2	<b>43.2(10)</b>	50.0(10)	55.4(10)	54.1(10)	58.1(10)
	3	40.5(10)	47.3(10)	48.6(10)	51.4(10)	56.8(10)
AC+T	1	2.7(10)	5.4(10)	5.4(10)	5.4(10)	5.4(10)
	2	5.4(10)	5.4(10)	5.4(10)	5.4(10)	5.4(10)
	3	5.4(10)	5.4(10)	5.4(10)	6.8(10)	8.1(10)
k10	1	45.9(10)	<b>59.5(10)*</b>	<b>74.3(10)</b>	<b>85.1(10)*</b>	<b>100.0(10)</b>
	2	41.9(10)*	<b>58.1(10)*</b>	<b>73.0(10)*</b>	<b>75.7(10)</b>	<b>100.0(10)*</b>
	3	35.1(10)*	<b>55.4(10)*</b>	<b>66.2(10)*</b>	<b>71.6(10)*</b>	<b>97.3(10)</b>
k10+T	1	5.4(10)	5.4(10)	5.4(10)	5.4(10)	10.8(10)
	2	5.4(10)	5.4(10)	5.4(10)	5.4(10)	10.8(10)
	3	4.1(10)	5.4(10)	5.4(10)	6.8(10)	16.2(10)
SEM	1	40.5(10)	55.4(10)	70.3(10)	81.1(10)	100.0(10)
	2	27.0(10)	50.0(10)	62.2(10)	67.6(10)	97.3(10)
	3	9.5(10)	31.1(10)	48.6(10)	59.5(10)	91.9(10)
SEM+T	1	44.6(10)	52.7(10)	66.2(10)	78.4(10)	97.3(10)
	2	27.0(10)	40.5(10)	52.7(10)	73.0(10)	81.1(10)
	3	17.6(10)	25.7(10)	41.9(10)	52.7(10)	75.7(9)
SEM.bic	1	41.9(10)	54.1(10)	70.3(10)	82.4(10)	100.0(10)
	2	33.8(10)	52.7(10)	67.6(10)	70.3(10)	98.6(10)
	3	9.5(10)	41.9(10)	51.4(10)	64.9(10)	91.9(10)
SEM.bic+T	1	40.5(10)	48.6(10)	63.5(10)	81.1(10)	98.6(10)
	2	29.7(10)	40.5(10)	56.8(10)	71.6(10)	91.9(9)
	3	16.2(10)	36.5(10)	39.2(10)	55.4(10)	81.1(9)
GES-EM	1	45.9(10)	56.8(10)	70.3(10)	78.4(10)	86.5(10)
	2	40.5(10)	51.4(10)	64.9(10)	70.3(10)	81.1(9)
	3	17.6(10)	41.9(10)	62.2(10)	64.9(10)	51.4(10)
GES-EM+T	1	18.9(10)	35.1(9)	31.1(10)	47.3(10)	64.9(10)
	2	18.9(10)	24.3(10)	29.7(10)	36.5(6)	56.8(8)
	3	-	-	-	-	48.6(3)
SEM+RT	1	40.5(10)	50.0(10)	67.6(10)	78.4(10)	91.9(10)
	2	31.1(10)	45.9(10)	56.8(10)	63.5(10)	90.5(10)†
	3	14.9(10)	37.8(10)†	44.6(10)	48.6(10)	78.4(10)

Recall Child mcar Recall in percentuale

Recall Insurance mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	35.0(10)	42.1(10)	45.0(10)	47.9(10)	53.6(10)
MWST-EM	1	17.9(10)	18.6(10)	20.7(10)	22.1(10)	25.7(10)
	2	17.1(10)	21.4(10)	21.4(10)	20.7(10)	24.3(10)
	3	20.7(10)	21.4(10)	22.1(10)	22.9(10)	24.3(10)
CC	1	25.0(10)	25.0(10)	40.0(10)	<b>45.7(10)</b>	47.1(10)
	2	<b>32.9(5)</b>	32.9(7)	17.9(10)	14.3(10)	30.0(10)
	3	-	30.0(3)	30.0(1)	25.7(1)	25.7(9)
CC+T	1	6.4(10)	7.1(10)	10.7(10)	11.4(10)	17.9(10)
	2	10.0(6)	7.1(10)	7.1(10)	4.3(10)	8.6(10)
	3	-	15.7(3)	15.7(1)	17.1(1)	11.4(10)
AC	1	<b>33.6(10)</b>	33.6(10)	42.9(10)	37.1(10)	40.7(10)
	2	27.1(10)	28.6(10)	35.7(10)	38.6(10)	44.3(10)
	3	<b>28.6(10)</b>	28.6(10)	32.1(10)	30.7(10)	43.6(10)
AC+T	1	14.3(10)	12.9(10)	15.0(10)	17.9(10)	15.7(10)
	2	15.0(10)	17.9(10)	15.7(10)	15.7(10)	15.7(10)
	3	17.9(10)	17.9(10)	17.9(10)	17.9(10)	17.9(10)
k10	1	30.7(10)	<b>37.1(10)</b>	<b>45.7(10)*</b>	44.3(10)*	<b>56.4(10)</b>
	2	27.1(10)*	<b>33.6(10)</b>	<b>37.1(10)</b>	<b>43.6(10)*</b>	<b>53.6(10)*</b>
	3	22.9(10)*	<b>31.4(10)*</b>	<b>34.3(10)*</b>	<b>40.7(10)*</b>	<b>49.3(10)</b>
k10+T	1	9.3(10)	10.0(10)	12.9(10)	14.3(10)	17.9(10)
	2	7.9(10)	9.3(10)	14.3(10)	14.3(10)	18.6(10)
	3	7.1(10)	10.0(10)	12.1(10)	14.3(10)	17.1(10)
SEM	1	27.1(10)	34.3(10)	37.1(10)	40.0(10)	53.6(10)
	2	18.6(10)	29.3(10)	37.1(10)	37.9(10)	52.9(9)
	3	8.6(10)	22.9(10)	30.7(10)	37.1(10)	43.6(4)
SEM+T	1	27.1(10)	33.6(10)	35.7(10)	40.7(10)	52.9(10)
	2	21.4(10)	29.3(10)	31.4(10)	31.4(10)	45.7(7)
	3	14.3(10)†	27.1(10)	30.0(10)	30.7(10)	44.3(1)
SEM.bic	1	22.9(10)	32.1(10)	40.7(10)	37.1(10)	50.7(10)
	2	17.1(10)	28.6(10)	35.7(10)	36.4(10)	47.1(9)
	3	9.3(10)	22.9(10)	32.9(10)	37.9(10)	40.0(6)
SEM.bic+T	1	27.1(10)	33.6(10)	35.0(10)	32.9(10)	55.0(10)
	2	18.6(10)	28.6(10)	32.9(10)	33.6(10)	48.6(8)
	3	14.3(10)	25.0(10)	30.0(10)	27.9(10)	37.1(2)
GES-EM	1	26.4(10)	34.3(10)	41.4(9)	38.6(9)	-
	2	22.1(10)	28.6(10)	35.7(9)	35.7(9)	-
	3	11.4(9)	21.4(10)	25.7(10)	26.4(10)	-
GES-EM+T	1	5.0(10)	6.4(10)	24.3(10)	27.1(10)	-
	2	12.9(10)	22.9(10)	26.4(10)	27.1(10)	-
	3	7.9(10)	17.1(10)	20.7(10)	21.4(8)	-
SEM+RT	1	26.4(10)	32.1(10)	35.7(10)	38.6(10)	52.9(10)
	2	18.6(10)	27.9(10)	32.9(10)	32.1(10)	44.3(7)
	3	12.9(10)	26.4(10)	28.6(10)	27.1(9)	40.0(2)

Recall Insurance mcar Recall in percentuale

Recall Alarm1 mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	56.0(10)	66.0(10)	81.0(10)	85.0(10)	92.0(10)
MWST-EM	1	32.0(10)	32.0(10)	32.0(10)	32.0(10)	32.0(10)
	2	32.0(10)	32.0(10)	30.0(10)	32.0(10)	32.0(10)
	3	26.0(10)	30.0(10)	32.0(10)	32.0(10)	32.0(10)
CC	1	47.0(10)	44.0(10)	48.0(10)	64.0(10)	73.0(10)
	2	6.0(3)	8.0(4)	32.0(2)	14.0(5)	47.0(10)
	3	-	-	8.0(1)	16.0(1)	0.0(1)
CC+T	1	17.0(10)	18.0(10)	22.0(10)	26.0(10)	26.0(10)
	2	29.0(4)	22.0(6)	20.0(3)	21.0(10)	16.0(10)
	3	-	-	24.0(1)	22.0(1)	24.0(1)
AC	1	45.0(10)	46.0(10)	55.0(10)	69.0(10)	69.0(10)
	2	<b>42.0(10)</b>	45.0(10)	53.0(10)	69.0(10)	70.0(10)
	3	<b>38.0(10)</b>	39.0(10)	51.0(10)	52.0(10)	70.0(10)
AC+T	1	25.0(10)	28.0(10)	28.0(10)	27.0(10)	28.0(10)
	2	26.0(10)	27.0(10)	26.0(10)	28.0(10)	26.0(10)
	3	25.0(10)	28.0(10)	28.0(10)	28.0(10)	27.0(10)
k10	1	52.0(10)	58.0(10)	72.0(10)	<b>80.0(10)</b>	87.0(10)
	2	42.0(10)	<b>53.0(10)</b>	62.0(10)	68.0(10)	<b>87.0(10)*</b>
	3	32.0(10)	39.0(10)	<b>53.0(10)</b>	59.0(10)	<b>77.0(10)</b>
k10+T	1	24.0(10)	25.0(10)	30.0(10)	29.0(10)	21.0(10)
	2	22.0(10)	24.0(10)	27.0(10)	27.0(10)	22.0(10)
	3	20.0(10)	24.0(10)	27.0(10)	30.0(10)	23.0(10)
SEM	1	43.0(10)	58.0(10)	69.0(10)	76.0(10)	83.0(10)
	2	38.0(10)	43.0(10)	58.0(10)	64.0(9)	76.0(8)
	3	21.0(10)	39.0(10)	45.0(10)	<b>65.0(8)</b>	68.0(1)
SEM+T	1	<b>60.0(10)*†</b>	<b>65.0(10)†</b>	<b>84.0(10)*†</b>	80.0(10)†	<b>94.0(10)†</b>
	2	41.0(10)	53.0(10)†	<b>65.0(10)†</b>	<b>70.0(10)†</b>	56.0(8)
	3	34.0(10)†	39.0(10)	50.0(10)†	62.0(9)†	76.0(5)
SEM.bic	1	42.0(10)	43.0(10)	62.0(10)	68.0(10)	67.0(10)
	2	36.0(10)	44.0(10)	58.0(10)	56.0(10)	74.0(7)
	3	20.0(10)	33.0(10)	39.0(10)	63.0(10)	40.0(3)
SEM.bic+T	1	41.0(10)	49.0(10)	68.0(10)	74.0(10)	77.0(10)
	2	38.0(10)	48.0(10)	59.0(10)	69.0(10)	58.0(10)
	3	28.0(10)	<b>43.0(10)</b>	53.0(10)	58.0(9)	66.0(6)
GES-EM	1	48.0(10)	48.0(9)	53.0(10)	-	-
	2	41.0(10)	50.0(10)	51.0(10)	53.0(6)	-
	3	22.0(10)	33.0(10)	37.0(10)	49.0(10)	-
GES-EM+T	1	5.0(10)	13.0(10)	42.0(10)	42.0(7)	-
	2	16.0(10)	33.0(10)	-	52.0(3)	-
	3	19.0(10)	33.0(10)	40.0(9)	-	-
SEM+RT	1	46.0(10)	52.0(10)	58.0(10)	69.0(10)	68.0(9)
	2	37.0(10)	43.0(10)	55.0(10)	52.0(9)	54.0(5)
	3	21.0(10)	35.0(10)	38.0(10)	44.0(7)	-

Recall Alarm1 mcar Recall in percentuale

Recall HailFinder mcar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	6.6(10)	7.2(10)	7.2(10)	8.4(10)	8.4(10)
MWST-EM	1	<b>12.0(10)</b>	<b>11.4(10)</b>	<b>14.5(10)</b>	<b>15.1(10)</b>	<b>26.5(10)</b>
	2	<b>10.2(10)</b>	<b>10.8(10)</b>	<b>12.0(10)</b>	<b>15.7(10)</b>	<b>16.3(10)</b>
	3	<b>9.6(10)</b>	<b>13.3(10)</b>	<b>13.9(10)</b>	<b>16.3(10)</b>	<b>16.9(10)</b>
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	6.0(10)	7.2(10)*	9.0(10)	7.8(10)	10.8(10)
	2	6.0(10)	7.2(10)	7.8(10)	10.2(10)	9.0(10)
	3	4.8(10)	6.0(10)	7.8(10)	7.2(10)	10.2(10)
k10+T	1	6.0(10)	6.6(10)	10.8(10)	10.8(10)	12.0(10)
	2	6.0(10)	6.0(10)	9.0(10)	10.8(10)	12.0(10)
	3	4.8(10)	6.0(10)	9.0(10)	10.8(10)	14.5(10)
SEM	1	3.6(10)	6.0(10)	9.6(10)	8.4(10)	10.8(8)
	2	2.4(10)	4.8(10)	8.4(10)	9.0(8)	6.0(1)
	3	0.0(10)	2.4(10)	4.8(10)	9.6(9)	-
SEM+T	1	4.8(10)†	6.0(10)	9.6(10)	10.8(10)*†	12.0(8)
	2	4.8(10)†	6.0(10)†	9.6(10)†	8.4(8)	9.6(2)
	3	4.8(10)†	6.0(10)†	8.4(10)†	9.6(10)*†	-
SEM.bic	1	4.8(10)	6.0(10)	9.6(10)	8.4(10)	12.0(5)
	2	3.6(10)	4.8(10)	8.4(10)	10.2(6)	-
	3	0.0(10)	3.6(10)	6.0(10)	8.4(7)	-
SEM.bic+T	1	6.0(10)	7.2(10)	9.0(10)	10.8(10)	12.0(6)
	2	4.8(10)	6.0(10)	8.4(10)	9.6(7)	-
	3	5.4(10)	6.6(10)	9.6(10)	9.6(9)	-
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	3.6(10)	5.4(10)	9.0(10)	8.4(10)	10.8(8)
	2	2.4(10)	5.4(10)	7.2(10)	10.2(8)	-
	3	0.0(10)	2.4(10)	4.8(9)	8.4(8)	-

**Recall HailFinder mcar** Recall in percentuale

Recall Hepar2 mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	10.8(10)	15.8(10)	26.2(10)	32.7(10)	48.1(10)
MWST-EM	1	<b>11.5(10)</b>	13.1(10)	11.9(10)	13.1(10)	13.1(10)
	2	<b>11.5(10)</b>	13.1(10)	13.1(10)	13.1(10)	13.1(10)
	3	<b>11.5(10)</b>	<b>13.1(10)</b>	13.1(10)	13.1(10)	13.1(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	10.4(10)*	<b>15.4(10)*</b>	<b>26.2(10)*</b>	<b>31.5(10)</b>	48.5(10)
	2	8.5(10)*	<b>13.8(10)*</b>	<b>21.2(10)*</b>	<b>27.7(10)</b>	45.0(10)
	3	7.3(10)*	10.4(10)*	<b>16.9(10)*</b>	<b>23.1(10)</b>	<b>42.3(10)</b>
k10+T	1	4.6(10)	6.2(10)	6.9(10)	8.1(10)	10.4(10)
	2	4.6(10)	5.8(10)	6.9(10)	9.2(10)	10.4(10)
	3	3.8(10)	4.6(10)	6.5(10)	7.7(10)	10.8(10)
SEM	1	6.9(10)	11.2(10)	22.3(10)	29.6(10)	<b>48.8(10)</b>
	2	5.0(10)	8.5(10)	16.2(10)	26.2(10)	<b>45.8(10)</b>
	3	1.2(10)	4.6(10)	10.8(10)	20.4(10)	39.2(10)
SEM+T	1	7.3(10)	11.5(10)	21.9(10)	29.2(10)	45.4(10)
	2	5.4(10)	10.4(10)†	16.5(10)†	26.5(10)	44.2(10)
	3	3.8(10)†	7.7(10)†	13.1(10)†	21.9(10)†	41.2(10)
SEM.bic	1	6.5(10)	11.5(10)	21.5(10)	30.4(10)	46.9(10)
	2	5.0(10)	7.7(10)	15.0(10)	25.8(10)	43.8(10)
	3	1.2(10)	5.4(10)	10.4(10)	20.0(10)	40.0(10)
SEM.bic+T	1	6.9(10)	11.5(10)	21.9(10)	30.0(10)	44.6(10)
	2	5.4(10)	9.6(10)	15.8(10)	27.3(10)	41.9(10)
	3	3.8(10)	7.7(10)	13.1(10)	21.2(10)	39.6(10)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	6.9(10)	10.4(10)	22.3(10)	28.5(10)	48.1(10)
	2	5.0(10)	8.5(10)	15.4(10)	26.2(10)	45.0(10)
	3	2.3(10)	5.0(10)	11.2(10)	20.0(10)	40.0(10)

Recall Hepar2 mcar Recall in percentuale

Recall Asia mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	36.4(10)	45.5(10)	45.5(10)	45.5(10)	72.7(10)
MWST-EM	1	<b>45.5(10)</b>	45.5(10)	45.5(10)	45.5(10)	45.5(10)
	2	<b>45.5(10)</b>	45.5(10)	45.5(10)	45.5(10)	45.5(10)
	3	<b>45.5(10)</b>	<b>45.5(10)</b>	45.5(10)	45.5(10)	45.5(10)
CC	1	27.3(10)	31.8(10)	45.5(10)	45.5(10)	63.6(10)
	2	27.3(10)	36.4(10)	45.5(10)	45.5(10)	45.5(10)
	3	36.4(10)	22.7(10)	45.5(10)	45.5(10)	45.5(10)
CC+T	1	27.3(10)	31.8(10)	27.3(10)	27.3(10)	27.3(10)
	2	27.3(10)	36.4(10)	27.3(10)	27.3(10)	31.8(10)
	3	36.4(10)	31.8(10)	31.8(10)	27.3(10)	27.3(10)
AC	1	27.3(10)	36.4(10)	45.5(10)	45.5(10)	54.5(10)
	2	31.8(10)	36.4(10)	45.5(10)	45.5(10)	63.6(10)
	3	27.3(10)	22.7(10)	45.5(10)	45.5(10)	45.5(10)
AC+T	1	31.8(10)	36.4(10)	40.9(10)	36.4(10)	36.4(10)
	2	36.4(10)	31.8(10)	40.9(10)	36.4(10)	36.4(10)
	3	40.9(10)	36.4(10)	36.4(10)	36.4(10)	36.4(10)
k10	1	31.8(10)	45.5(10)	59.1(10)	45.5(10)	77.3(10)
	2	36.4(10)	45.5(10)	45.5(10)	50.0(10)	77.3(10)
	3	45.5(10)	45.5(10)	45.5(10)	45.5(10)	68.2(10)
k10+T	1	36.4(10)	27.3(10)	31.8(10)	27.3(10)	27.3(10)
	2	36.4(10)	27.3(10)	27.3(10)	27.3(10)	27.3(10)
	3	40.9(10)	27.3(10)	27.3(10)	27.3(10)	27.3(10)
SEM	1	36.4(10)	50.0(10)	72.7(10)	<b>81.8(10)</b>	<b>81.8(10)</b>
	2	31.8(10)	40.9(10)	54.5(10)	<b>77.3(10)</b>	68.2(10)
	3	27.3(10)	36.4(10)	54.5(10)	59.1(10)	77.3(10)
SEM+T	1	36.4(10)	50.0(10)*	<b>77.3(10)*</b>	77.3(10)*	81.8(10)
	2	36.4(10)	40.9(10)	<b>77.3(10)*</b>	72.7(10)*	<b>81.8(10)</b>
	3	27.3(10)	45.5(10)	<b>63.6(10)*</b>	<b>68.2(10)*</b>	77.3(10)
SEM.bic	1	36.4(10)	<b>54.5(10)</b>	72.7(10)	81.8(10)	72.7(10)
	2	27.3(10)	45.5(10)	54.5(10)	77.3(10)	81.8(10)
	3	22.7(10)	45.5(10)	54.5(10)	63.6(10)	77.3(10)
SEM.bic+T	1	31.8(10)	45.5(10)	72.7(10)	81.8(10)	81.8(10)
	2	40.9(10)	45.5(10)	68.2(10)	72.7(10)	81.8(10)
	3	36.4(10)	45.5(10)	59.1(10)	63.6(10)	<b>81.8(10)</b>
GES-EM	1	40.9(10)	54.5(10)	72.7(10)	77.3(10)	81.8(10)
	2	36.4(10)	<b>54.5(10)</b>	54.5(9)	77.3(10)	81.8(9)
	3	36.4(10)	45.5(10)	63.6(10)	59.1(10)	77.3(10)
GES-EM+T	1	45.5(9)	45.5(10)	45.5(10)	45.5(10)	81.8(10)
	2	40.9(10)	40.9(10)	50.0(10)	45.5(10)	81.8(9)
	3	36.4(9)	36.4(10)	45.5(10)	50.0(10)	81.8(10)
SEM+RT	1	36.4(10)	50.0(10)	81.8(10)	72.7(10)	81.8(10)
	2	36.4(10)	50.0(10)	54.5(10)	72.7(10)	81.8(10)
	3	27.3(10)	36.4(10)	63.6(10)	54.5(10)	63.6(10)

Recall Asia mar Recall in percentuale

Recall Child mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	45.9(10)	66.2(10)	77.0(10)	89.2(10)	100.0(10)
MWST-EM	1	5.4(10)	5.4(10)	5.4(10)	8.1(10)	10.8(10)
	2	6.8(10)	5.4(10)	5.4(10)	10.8(10)	13.5(10)
	3	5.4(10)	8.1(10)	6.8(10)	8.1(10)	10.8(10)
CC	1	24.3(10)	33.8(10)	50.0(10)	73.0(10)	83.8(10)
	2	27.0(10)	28.4(10)	33.8(10)	50.0(10)	75.7(10)
	3	40.5(9)	24.3(10)	27.0(10)	28.4(10)	56.8(10)
CC+T	1	2.7(10)	5.4(10)	5.4(10)	5.4(10)	8.1(10)
	2	2.7(10)	4.1(10)	5.4(10)	5.4(10)	8.1(10)
	3	5.4(9)	2.7(10)	2.7(10)	4.1(10)	5.4(10)
AC	1	<b>45.9(10)</b>	45.9(10)	52.7(10)	68.9(10)	62.2(10)
	2	35.1(10)	47.3(10)	58.1(10)	55.4(10)	60.8(10)
	3	<b>41.9(10)</b>	44.6(10)	50.0(10)	56.8(10)	55.4(10)
AC+T	1	5.4(10)	5.4(10)	5.4(10)	6.8(10)	10.8(10)
	2	5.4(10)	5.4(10)	4.1(10)	5.4(10)	8.1(10)
	3	5.4(10)	5.4(10)	5.4(10)	5.4(10)	9.5(10)
k10	1	45.9(10)*	<b>62.2(10)*</b>	<b>74.3(10)</b>	<b>83.8(10)*</b>	<b>100.0(10)*</b>
	2	<b>40.5(10)</b>	<b>59.5(10)*</b>	<b>66.2(10)*</b>	<b>77.0(10)</b>	<b>100.0(10)*</b>
	3	35.1(10)*	<b>60.8(10)*</b>	<b>66.2(10)</b>	<b>75.7(10)*</b>	<b>97.3(10)*</b>
k10+T	1	5.4(10)	5.4(10)	5.4(10)	5.4(10)	10.8(10)
	2	5.4(10)	5.4(10)	5.4(10)	10.8(10)	13.5(10)
	3	5.4(10)	5.4(10)	5.4(10)	5.4(10)	10.8(10)
SEM	1	33.8(10)	55.4(10)	63.5(10)	78.4(10)	95.9(10)
	2	29.7(10)	51.4(10)	55.4(10)	71.6(10)	93.2(10)
	3	24.3(10)	37.8(10)	48.6(10)	59.5(10)	89.2(10)
SEM+T	1	33.8(10)	52.7(10)	64.9(10)	71.6(10)	95.9(10)
	2	33.8(10)	45.9(10)	59.5(10)	68.9(10)	82.4(10)
	3	28.4(10)	41.9(10)	55.4(10)†	51.4(10)	70.3(10)
SEM.bic	1	33.8(10)	54.1(10)	66.2(10)	78.4(10)	94.6(10)
	2	33.8(10)	48.6(10)	60.8(10)	75.7(10)	94.6(10)
	3	25.7(10)	40.5(10)	55.4(10)	67.6(10)	85.1(10)
SEM.bic+T	1	33.8(10)	44.6(10)	67.6(10)	63.5(10)	91.9(10)
	2	31.1(10)	47.3(10)	58.1(10)	63.5(10)	91.9(10)
	3	27.0(10)	40.5(10)	52.7(10)	59.5(10)	85.1(10)
GES-EM	1	36.5(10)	59.5(10)	70.3(10)	77.0(10)	58.1(8)
	2	35.1(10)	54.1(10)	63.5(10)	70.3(10)	83.8(10)
	3	25.7(10)	48.6(10)	59.5(10)	68.9(10)	81.1(9)
GES-EM+T	1	16.2(10)	24.3(8)	32.4(9)	31.1(10)	45.9(9)
	2	12.2(10)	16.2(10)	27.0(10)	32.4(7)	60.8(6)
	3	-	-	-	-	37.8(7)
SEM+RT	1	35.1(10)	54.1(10)	63.5(10)	67.6(10)	93.2(10)
	2	31.1(10)	50.0(10)	52.7(10)	55.4(10)	91.9(10)
	3	31.1(10)	44.6(10)	48.6(10)	62.2(10)	77.0(10)

Recall Child mar Recall in percentuale

Recall Insurance mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>35.0(10)</i>	<i>42.1(10)</i>	<i>45.0(10)</i>	<i>47.9(10)</i>	<i>53.6(10)</i>
MWST-EM	1	20.0(10)	20.7(10)	20.0(10)	22.9(10)	24.3(10)
	2	18.6(10)	20.7(10)	20.0(10)	20.0(10)	23.6(10)
	3	18.6(10)	22.9(10)	21.4(10)	22.9(10)	23.6(10)
CC	1	17.9(10)	22.9(10)	27.9(10)	36.4(10)	47.1(10)
	2	16.4(4)	20.0(9)	19.3(10)	24.3(9)	40.0(10)
	3	<b>27.9(4)</b>	25.0(8)	19.3(8)	20.0(6)	25.0(10)
CC+T	1	3.6(10)	6.4(10)	9.3(10)	12.9(10)	15.7(10)
	2	4.3(9)	4.3(10)	4.3(10)	7.9(10)	12.9(10)
	3	8.6(5)	2.9(9)	5.7(10)	3.6(10)	8.6(10)
AC	1	27.1(10)	31.4(10)	37.1(10)	42.1(10)	48.6(10)
	2	<b>27.1(10)</b>	31.4(10)	35.0(10)	37.9(10)	48.6(10)
	3	25.7(10)	31.4(10)	36.4(10)	40.7(10)	43.6(10)
AC+T	1	12.1(10)	13.6(10)	14.3(10)	17.9(10)	17.1(10)
	2	12.9(10)	13.6(10)	12.9(10)	14.3(10)	16.4(10)
	3	10.7(10)	15.0(10)	14.3(10)	15.0(10)	17.1(10)
k10	1	<b>32.9(10)</b>	<b>37.1(10)</b>	<b>44.3(10)*</b>	<b>46.4(10)</b>	<b>55.7(10)</b>
	2	27.1(10)	<b>34.3(10)*</b>	<b>42.9(10)*</b>	<b>46.4(10)</b>	<b>51.4(10)</b>
	3	25.7(10)	<b>33.6(10)*</b>	<b>41.4(10)*</b>	<b>41.4(10)*</b>	<b>49.3(10)</b>
k10+T	1	8.6(10)	11.4(10)	12.1(10)	15.7(10)	17.1(10)
	2	8.6(10)	9.3(10)	13.6(10)	15.7(10)	14.3(10)
	3	8.6(10)	7.9(10)	12.1(10)	15.0(10)	15.7(10)
SEM	1	25.7(10)	28.6(10)	36.4(10)	42.1(10)	46.4(10)
	2	22.9(10)	28.6(10)	35.0(10)	39.3(10)	46.4(10)
	3	17.9(10)	26.4(10)	32.9(10)	35.0(10)	40.0(10)
SEM+T	1	27.1(10)	32.1(10)	34.3(10)	45.0(10)	52.9(10)
	2	24.3(10)	30.7(10)	30.7(10)	36.4(10)	46.4(10)
	3	23.6(10)†	28.6(10)†	32.1(10)	34.3(10)	45.7(10)
SEM.bic	1	25.7(10)	31.4(10)	36.4(10)	42.9(10)	51.4(10)
	2	24.3(10)	29.3(10)	37.9(10)	36.4(10)	42.9(10)
	3	19.3(10)	25.0(10)	32.9(10)	35.0(10)	44.3(10)
SEM.bic+T	1	28.6(10)	31.4(10)	35.7(10)	39.3(10)	48.6(10)
	2	27.1(10)	30.7(10)	30.0(10)	38.6(10)	38.6(10)
	3	23.6(10)	28.6(10)	33.6(10)	35.0(10)	43.6(10)
GES-EM	1	30.0(10)	33.6(10)	39.3(10)	42.9(10)	-
	2	25.7(10)	30.0(10)	37.1(9)	38.6(7)	-
	3	22.9(10)	24.3(10)	33.6(10)	35.0(8)	-
GES-EM+T	1	7.1(9)	1.4(9)	10.0(10)	17.9(10)	-
	2	7.1(9)	5.0(10)	8.6(10)	14.3(10)	-
	3	10.0(9)	10.7(10)	10.0(10)	11.4(9)	-
SEM+RT	1	<i>25.7(10)</i>	<i>31.4(10)</i>	<i>33.6(10)</i>	<i>38.6(10)</i>	<i>45.0(10)</i>
	2	<i>20.7(10)</i>	<i>31.4(10)</i>	<i>30.7(10)</i>	<i>35.7(10)</i>	<i>45.7(10)</i>
	3	<i>22.9(10)</i>	<i>25.7(10)</i>	<i>32.1(10)</i>	<i>31.4(10)</i>	<i>40.7(10)</i>

Recall Insurance mar Recall in percentuale

Recall Alarm1 mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>56.0(10)</i>	<i>66.0(10)</i>	<i>81.0(10)</i>	<i>85.0(10)</i>	<i>92.0(10)</i>
MWST-EM	1	32.0(10)	32.0(10)	32.0(10)	32.0(10)	32.0(10)
	2	29.0(10)	32.0(10)	32.0(10)	32.0(10)	32.0(10)
	3	32.0(10)	30.0(10)	32.0(10)	32.0(10)	32.0(10)
CC	1	<b>55.0(6)</b>	51.0(10)	44.0(10)	42.0(10)	67.0(10)
	2	7.0(4)	28.0(2)	45.0(10)	47.0(6)	47.0(10)
	3	6.0(2)	6.0(3)	50.0(1)	27.0(6)	32.0(7)
CC+T	1	16.0(10)	18.0(10)	20.0(10)	22.0(10)	26.0(10)
	2	24.0(5)	22.0(7)	15.0(10)	18.0(10)	21.0(10)
	3	26.0(3)	22.0(4)	16.0(1)	20.0(7)	21.0(8)
AC	1	36.0(10)	40.0(10)	58.0(10)	69.0(10)	69.0(10)
	2	37.0(10)	40.0(10)	56.0(10)	59.0(10)	66.0(10)
	3	34.0(10)	43.0(10)	42.0(10)	53.0(10)	67.0(10)
AC+T	1	25.0(10)	27.0(10)	29.0(10)	28.0(10)	25.0(10)
	2	24.0(10)	27.0(10)	28.0(10)	27.0(10)	23.0(10)
	3	26.0(10)	24.0(10)	28.0(10)	26.0(10)	25.0(10)
k10	1	45.0(10)	60.0(10)	70.0(10)	73.0(10)	89.0(10)
	2	47.0(10)	46.0(10)	<b>67.0(10)</b>	<b>76.0(10)</b>	<b>93.0(10)</b>
	3	<b>44.0(10)</b>	50.0(10)	<b>68.0(10)</b>	70.0(10)	<b>80.0(10)</b>
k10+T	1	23.0(10)	27.0(10)	26.0(10)	26.0(10)	22.0(10)
	2	20.0(10)	26.0(10)	28.0(10)	28.0(10)	20.0(10)
	3	23.0(10)	24.0(10)	27.0(10)	26.0(10)	23.0(10)
SEM	1	44.0(10)	41.0(10)	62.0(10)	69.0(10)	71.0(10)
	2	30.0(10)	38.0(10)	49.0(10)	54.0(10)	70.0(10)
	3	31.0(10)	38.0(10)	46.0(10)	60.0(10)	63.0(10)
SEM+T	1	55.0(10)†	<b>66.0(10)*†</b>	<b>82.0(10)†</b>	<b>79.0(10)†</b>	<b>94.0(10)†</b>
	2	<b>48.0(10)†</b>	<b>58.0(9)*†</b>	66.0(10)†	63.0(10)†	88.0(10)†
	3	44.0(10)†	<b>58.0(9)†</b>	66.0(10)†	<b>73.0(10)†</b>	65.0(10)
SEM.bic	1	40.0(10)	44.0(10)	48.0(10)	63.0(10)	59.0(10)
	2	32.0(10)	36.0(10)	52.0(10)	57.0(10)	58.0(10)
	3	28.0(10)	36.0(10)	45.0(10)	55.0(10)	53.0(10)
SEM.bic+T	1	39.0(10)	50.0(10)	64.0(10)	62.0(10)	81.0(10)
	2	34.0(10)	45.0(10)	58.0(10)	64.0(10)	81.0(10)
	3	34.0(10)	47.0(10)	57.0(10)	66.0(10)	64.0(10)
GES-EM	1	48.0(10)	36.0(10)	48.0(9)	50.0(3)	-
	2	33.0(10)	34.0(9)	46.0(9)	48.0(3)	-
	3	33.0(10)	34.0(10)	32.0(9)	56.0(1)	-
GES-EM+T	1	4.0(10)	8.0(10)	16.0(10)	12.0(9)	-
	2	4.0(10)	12.0(10)	20.0(10)	18.0(10)	-
	3	4.0(10)	12.0(10)	16.0(10)	12.0(9)	-
SEM+RT	1	<i>41.0(10)</i>	<i>40.0(10)</i>	<i>57.0(10)</i>	<i>61.0(10)</i>	<i>72.0(10)</i>
	2	<i>37.0(10)</i>	<i>40.0(9)</i>	<i>55.0(10)</i>	<i>54.0(10)</i>	<i>59.0(10)</i>
	3	<i>25.0(10)</i>	<i>34.0(10)</i>	<i>50.0(10)</i>	<i>51.0(10)</i>	<i>55.0(10)</i>

Recall Alarm1 mar Recall in percentuale

Recall HailFinder mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	6.6(10)	7.2(10)	7.2(10)	8.4(10)	8.4(10)
MWST-EM	1	<b>13.3(10)</b>	<b>12.0(10)</b>	<b>13.3(10)</b>	<b>14.5(10)</b>	<b>16.9(10)</b>
	2	<b>12.0(10)</b>	<b>10.8(10)</b>	<b>12.7(10)</b>	<b>14.5(10)</b>	<b>17.5(10)</b>
	3	<b>11.4(10)</b>	<b>13.3(10)</b>	<b>14.5(10)</b>	<b>16.9(10)</b>	<b>18.1(10)</b>
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	6.0(10)	7.2(10)*	7.2(10)	8.4(10)	8.4(10)
	2	6.0(10)	7.2(10)	8.4(10)	8.4(10)	9.6(10)
	3	4.8(10)	6.0(10)	7.2(10)	8.4(10)	9.6(10)
k10+T	1	6.0(10)	7.2(10)	8.4(10)	9.6(10)	12.0(10)
	2	5.4(10)	6.0(10)	9.6(10)	9.6(10)	10.8(10)
	3	4.8(10)	6.0(10)	10.2(10)	10.8(10)	12.0(10)
SEM	1	3.0(10)	4.8(10)	7.2(10)	8.4(10)	10.2(10)
	2	2.4(10)	4.8(10)	6.0(10)	7.8(10)	9.6(5)
	3	1.2(10)	3.6(10)	6.0(10)	7.2(10)	9.6(8)
SEM+T	1	4.8(10)†	6.0(10)†	8.4(10)	9.6(10)*†	9.6(10)*
	2	4.8(10)†	6.0(10)†	8.4(10)	10.8(9)†	9.6(9)
	3	4.8(10)†	6.0(10)†	9.0(10)†	10.8(10)†	12.0(7)†
SEM.bic	1	3.6(10)	4.8(10)	8.4(10)	8.4(10)	9.6(8)
	2	2.4(10)	4.8(10)	7.2(10)	8.4(10)	11.4(6)
	3	1.8(10)	4.2(10)	6.0(10)	7.2(10)	9.0(8)
SEM.bic+T	1	4.8(10)	6.0(10)	9.0(10)	9.6(10)	9.6(9)
	2	4.8(10)	6.6(10)	8.4(10)	9.6(10)	9.6(8)
	3	5.4(10)	7.2(10)	9.6(10)	9.6(10)	10.2(6)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	3.6(10)	4.8(10)	7.2(10)	8.4(10)	10.8(10)
	2	2.4(10)	4.8(10)	6.6(10)	7.2(10)	9.6(7)
	3	2.4(10)	3.6(10)	6.0(10)	7.2(10)	9.6(8)

Recall HailFinder mar Recall in percentuale

Recall Hepar2 mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	10.8(10)	15.8(10)	26.2(10)	32.7(10)	48.1(10)
MWST-EM	1	<b>11.5(10)</b>	11.9(10)	13.1(10)	13.1(10)	13.5(10)
	2	<b>11.5(10)</b>	12.3(10)	12.3(10)	13.1(10)	13.5(10)
	3	<b>11.5(10)</b>	<b>12.7(10)</b>	13.1(10)	13.1(10)	13.1(6)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	9.6(10)*	<b>13.1(10)*</b>	<b>23.5(10)</b>	<b>31.2(10)</b>	46.9(10)
	2	8.8(10)*	<b>14.6(10)*</b>	<b>21.9(10)</b>	<b>28.5(10)</b>	43.1(10)
	3	8.8(10)*	11.9(10)*	<b>22.7(10)*</b>	25.8(10)	39.2(10)
k10+T	1	4.6(10)	6.9(10)	6.9(10)	8.5(10)	10.8(10)
	2	5.0(10)	5.4(10)	6.9(10)	8.8(10)	11.5(10)
	3	4.6(10)	4.6(10)	6.9(10)	8.5(10)	10.0(6)
SEM	1	6.2(10)	10.4(10)	21.2(10)	30.4(10)	46.5(10)
	2	5.8(10)	10.0(10)	18.8(10)	28.1(10)	<b>44.6(10)</b>
	3	5.0(10)	8.1(10)	18.8(10)	<b>27.3(10)</b>	42.3(10)
SEM+T	1	6.9(10)	10.8(10)	21.2(10)	29.2(10)	<b>47.3(10)†</b>
	2	6.2(10)	10.8(10)	20.0(10)	28.1(10)	43.5(10)
	3	6.2(10)	10.4(10)†	18.8(10)	25.8(10)	<b>42.7(6)</b>
SEM.bic	1	6.2(10)	10.0(10)	21.9(10)	28.5(10)	45.4(10)
	2	5.4(10)	10.0(10)	18.1(10)	26.9(10)	43.1(10)
	3	4.6(10)	7.7(10)	18.5(10)	26.2(10)	40.0(10)
SEM.bic+T	1	6.9(10)	10.8(10)	21.2(10)	28.5(10)	44.2(10)
	2	5.4(10)	10.8(10)	19.2(10)	27.7(10)	40.8(10)
	3	5.8(10)	10.0(10)	18.8(10)	25.8(10)	42.7(6)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	6.2(10)	10.4(10)	20.0(10)	29.2(10)	44.6(10)
	2	5.8(10)	10.0(10)	19.6(10)	27.7(10)	43.1(10)
	3	5.4(10)	8.1(10)	18.8(10)	26.2(10)	41.2(10)

**Recall Hepar2 mar** Recall in percentuale

Recall Asia nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	36.4(10)	45.5(10)	45.5(10)	45.5(10)	72.7(10)
MWST-EM	1	<b>45.5(10)</b>	45.5(10)	45.5(10)	45.5(10)	45.5(10)
	2	<b>45.5(10)</b>	45.5(10)	45.5(10)	45.5(10)	45.5(10)
	3	<b>45.5(10)</b>	<b>45.5(10)</b>	45.5(10)	45.5(10)	45.5(10)
CC	1	36.4(10)	36.4(10)	45.5(10)	45.5(10)	63.6(10)
	2	36.4(10)	31.8(10)	45.5(10)	45.5(10)	63.6(10)
	3	27.3(10)	22.7(10)	31.8(10)	45.5(10)	45.5(10)
CC+T	1	45.5(10)	40.9(10)	27.3(10)	27.3(10)	40.9(10)
	2	36.4(10)	27.3(10)	27.3(10)	27.3(10)	27.3(10)
	3	40.9(10)	27.3(10)	27.3(10)	27.3(10)	36.4(10)
AC	1	36.4(10)	36.4(10)	45.5(10)	45.5(10)	63.6(10)
	2	27.3(10)	40.9(10)	45.5(10)	45.5(10)	63.6(10)
	3	18.2(10)	27.3(10)	45.5(10)	45.5(10)	63.6(10)
AC+T	1	36.4(10)	27.3(10)	31.8(10)	27.3(10)	36.4(10)
	2	40.9(10)	36.4(10)	36.4(10)	27.3(10)	36.4(10)
	3	36.4(10)	40.9(10)	36.4(10)	27.3(10)	36.4(10)
k10	1	27.3(10)	40.9(10)	50.0(10)	45.5(10)	63.6(10)
	2	36.4(10)	36.4(10)	45.5(10)	50.0(10)	63.6(10)
	3	31.8(10)	31.8(10)	45.5(10)	54.5(10)	63.6(10)
k10+T	1	40.9(10)	27.3(10)	27.3(10)	27.3(10)	27.3(10)
	2	40.9(10)	27.3(10)	27.3(10)	27.3(10)	27.3(10)
	3	27.3(10)	27.3(10)	27.3(10)	27.3(10)	27.3(10)
SEM	1	36.4(10)	45.5(10)	68.2(10)	72.7(10)	77.3(10)
	2	36.4(10)	36.4(10)	<b>72.7(10)</b>	72.7(10)	72.7(10)
	3	22.7(10)	36.4(10)	50.0(10)	54.5(10)	<b>81.8(10)</b>
SEM+T	1	31.8(10)	50.0(10)	63.6(10)	77.3(10)*	77.3(10)
	2	36.4(10)	40.9(10)	72.7(10)*	72.7(10)*	<b>81.8(10)*</b>
	3	36.4(10)†	40.9(10)*	50.0(10)	54.5(10)	77.3(10)
SEM.bic	1	40.9(10)	50.0(10)	<b>72.7(10)</b>	77.3(10)	<b>81.8(10)</b>
	2	31.8(10)	45.5(10)	72.7(10)	63.6(10)	77.3(10)
	3	18.2(10)	36.4(10)	59.1(10)	54.5(10)	81.8(10)
SEM.bic+T	1	45.5(10)	<b>54.5(10)</b>	72.7(10)	<b>81.8(10)</b>	81.8(10)
	2	40.9(10)	<b>50.0(10)</b>	72.7(10)	77.3(10)	81.8(10)
	3	27.3(10)	40.9(10)	54.5(10)	54.5(10)	72.7(10)
GES-EM	1	45.5(10)	45.5(9)	72.7(9)	81.8(10)	81.8(10)
	2	36.4(10)	45.5(10)	72.7(10)	<b>81.8(10)</b>	81.8(10)
	3	27.3(10)	40.9(10)	<b>63.6(10)</b>	<b>59.1(10)</b>	63.6(10)
GES-EM+T	1	36.4(10)	36.4(10)	31.8(10)	36.4(10)	81.8(10)
	2	36.4(7)	31.8(10)	45.5(10)	45.5(10)	54.5(10)
	3	22.7(10)	45.5(10)	40.9(10)	54.5(10)	59.1(10)
SEM+RT	1	40.9(10)	54.5(10)	54.5(10)	72.7(10)	72.7(10)
	2	27.3(10)	40.9(10)	72.7(10)	72.7(10)	81.8(10)
	3	18.2(10)	40.9(10)	45.5(10)	50.0(10)	72.7(10)

Recall Asia nmar Recall in percentuale

Recall Child nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	45.9(10)	66.2(10)	77.0(10)	89.2(10)	100.0(10)
MWST-EM	1	5.4(10)	5.4(10)	5.4(10)	5.4(10)	10.8(10)
	2	5.4(10)	5.4(10)	5.4(10)	10.8(10)	10.8(10)
	3	5.4(10)	8.1(10)	5.4(10)	8.1(10)	10.8(10)
CC	1	23.0(10)	27.0(10)	48.6(10)	74.3(10)	79.7(10)
	2	25.7(10)	24.3(10)	24.3(10)	45.9(10)	78.4(10)
	3	32.4(9)	29.7(10)	24.3(10)	27.0(10)	60.8(10)
CC+T	1	1.4(10)	5.4(10)	5.4(10)	5.4(10)	9.5(10)
	2	2.7(10)	4.1(10)	5.4(10)	5.4(10)	6.8(10)
	3	2.7(9)	5.4(10)	2.7(10)	5.4(10)	5.4(10)
AC	1	36.5(10)	43.2(10)	59.5(10)	<b>78.4(10)</b>	55.4(10)
	2	33.8(10)	41.9(10)	56.8(10)	56.8(10)	59.5(10)
	3	35.1(10)	37.8(10)	52.7(10)	55.4(10)	66.2(10)
AC+T	1	5.4(10)	5.4(10)	5.4(10)	5.4(10)	8.1(10)
	2	5.4(10)	5.4(10)	5.4(10)	8.1(10)	9.5(10)
	3	5.4(10)	8.1(10)	5.4(10)	5.4(10)	9.5(10)
k10	1	<b>48.6(10)*</b>	<b>62.2(10)*</b>	<b>73.0(10)*</b>	77.0(10)	<b>100.0(10)*</b>
	2	<b>48.6(10)*</b>	<b>59.5(10)*</b>	<b>70.3(10)*</b>	<b>71.6(10)</b>	<b>100.0(10)*</b>
	3	<b>44.6(10)*</b>	<b>59.5(10)*</b>	<b>64.9(10)*</b>	59.5(10)	<b>97.3(10)*</b>
k10+T	1	5.4(10)	5.4(10)	5.4(10)	5.4(10)	10.8(10)
	2	5.4(10)	5.4(10)	5.4(10)	8.1(10)	10.8(10)
	3	5.4(10)	5.4(10)	5.4(10)	8.1(10)	10.8(10)
SEM	1	28.4(10)	52.7(10)	66.2(10)	75.7(10)	93.2(10)
	2	31.1(10)	43.2(10)	54.1(10)	64.9(10)	93.2(10)
	3	27.0(10)	43.2(10)	52.7(10)	68.9(10)	91.9(10)
SEM+T	1	35.1(10)	45.9(10)	50.0(10)	68.9(10)	91.9(10)
	2	32.4(10)	43.2(10)	52.7(10)	66.2(10)	89.2(10)
	3	31.1(10)	44.6(10)	44.6(10)	64.9(10)	63.5(10)
SEM.bic	1	31.1(10)	55.4(10)	67.6(10)	75.7(10)	95.9(10)
	2	28.4(10)	48.6(10)	54.1(10)	64.9(10)	97.3(10)
	3	25.7(10)	44.6(10)	48.6(10)	<b>70.3(10)</b>	94.6(10)
SEM.bic+T	1	35.1(10)	40.5(10)	58.1(10)	67.6(10)	94.6(10)
	2	29.7(10)	40.5(10)	51.4(10)	71.6(10)	70.3(10)
	3	27.0(10)	39.2(10)	40.5(10)	63.5(10)	85.1(10)
GES-EM	1	36.5(10)	51.4(10)	67.6(10)	74.3(10)	81.1(9)
	2	32.4(10)	51.4(10)	67.6(10)	70.3(10)	79.7(8)
	3	29.7(10)	50.0(10)	58.1(10)	68.9(10)	79.7(8)
GES-EM+T	1	13.5(10)	25.7(10)	29.7(9)	37.8(10)	54.1(9)
	2	16.2(10)	18.9(9)	23.0(10)	29.7(4)	52.7(4)
	3	-	-	-	-	40.5(7)
SEM+RT	1	33.8(10)	48.6(10)	62.2(10)	70.3(10)	91.9(10)
	2	35.1(10)	47.3(10)	50.0(10)	58.1(10)	94.6(10)†
	3	29.7(10)	43.2(10)	48.6(10)	59.5(10)	81.1(10)

Recall Child nmar Recall in percentuale

Recall Insurance nmar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	35.0(10)	42.1(10)	45.0(10)	47.9(10)	53.6(10)
MWST-EM	1	17.9(10)	22.9(10)	20.7(10)	21.4(10)	26.4(10)
	2	17.1(10)	19.3(10)	21.4(10)	22.9(10)	24.3(10)
	3	17.1(10)	21.4(10)	21.4(10)	22.1(10)	24.3(10)
CC	1	15.7(10)	22.1(10)	29.3(10)	36.4(10)	45.0(10)
	2	24.3(7)	17.1(9)	22.1(10)	25.0(10)	42.9(10)
	3	25.0(4)	21.4(7)	20.0(8)	19.3(8)	22.9(10)
CC+T	1	4.3(10)	7.1(10)	8.6(10)	10.7(10)	17.1(10)
	2	4.3(10)	2.9(10)	5.0(10)	10.0(10)	12.1(10)
	3	5.7(9)	6.4(10)	5.7(10)	8.6(9)	7.9(10)
AC	1	<b>32.9(10)</b>	32.9(10)	<b>40.0(10)</b>	45.0(10)	52.9(10)
	2	<b>30.0(10)</b>	<b>37.1(10)</b>	34.3(10)	42.9(10)	47.1(10)
	3	25.0(10)	32.9(10)	34.3(10)	<b>42.9(10)</b>	44.3(10)
AC+T	1	11.4(10)	12.1(10)	12.9(10)	15.7(10)	17.1(10)
	2	11.4(10)	12.9(10)	15.7(10)	17.1(10)	17.1(10)
	3	11.4(10)	14.3(10)	14.3(10)	16.4(10)	18.6(10)
k10	1	31.4(10)*	<b>37.9(10)</b>	37.9(10)	<b>45.7(10)*</b>	<b>60.7(10)</b>
	2	30.0(10)	32.1(10)*	<b>37.9(10)</b>	<b>45.7(10)</b>	<b>55.0(10)</b>
	3	<b>27.9(10)</b>	<b>33.6(10)*</b>	35.0(10)	40.7(10)	<b>49.3(10)</b>
k10+T	1	10.0(10)	9.3(10)	14.3(10)	15.0(10)	18.6(10)
	2	9.3(10)	8.6(10)	13.6(10)	17.1(10)	16.4(10)
	3	9.3(10)	10.0(10)	12.9(10)	14.3(10)	17.1(10)
SEM	1	24.3(10)	31.4(10)	37.1(10)	37.1(10)	52.9(10)
	2	18.6(10)	30.0(10)	35.7(10)	38.6(10)	46.4(10)
	3	22.9(10)	27.1(10)	30.0(10)	35.7(10)	39.3(10)
SEM+T	1	25.7(10)	30.0(10)	37.9(10)	41.4(10)†	57.9(10)†
	2	22.9(10)	27.1(10)	36.4(10)	40.7(10)	47.1(8)†
	3	23.6(10)†	27.1(10)†	32.1(10)	32.9(10)	47.9(10)†
SEM.bic	1	24.3(10)	32.9(10)	35.7(10)	38.6(10)	44.3(10)
	2	20.0(10)	27.1(10)	35.7(10)	40.7(10)	46.4(10)
	3	17.9(10)	25.7(10)	<b>35.7(10)</b>	37.9(10)	45.0(10)
SEM.bic+T	1	24.3(10)	32.1(10)	34.3(10)	37.1(10)	44.3(10)
	2	24.3(10)	27.9(10)	31.4(10)	35.7(10)	40.7(10)
	3	21.4(10)	26.4(10)	35.0(10)	29.3(10)	46.4(10)
GES-EM	1	25.7(10)	32.1(10)	38.6(10)	34.3(7)	-
	2	22.9(10)	28.6(9)	34.3(9)	29.3(8)	-
	3	22.1(10)	27.9(10)	33.6(10)	35.0(10)	-
GES-EM+T	1	7.9(10)	1.4(9)	2.1(10)	7.9(10)	-
	2	7.1(10)	7.1(9)	3.6(10)	10.7(10)	-
	3	14.3(9)	9.3(10)	12.9(9)	19.3(10)	-
SEM+RT	1	27.9(10)	30.7(10)	35.0(10)	37.1(10)	45.0(10)
	2	22.1(10)	28.6(10)	35.0(10)	32.9(10)	40.7(10)
	3	20.7(10)	23.6(10)	30.0(10)	34.3(10)	41.4(10)

Recall Insurance nmar Recall in percentuale

Recall Alarm1 nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	56.0(10)	66.0(10)	81.0(10)	85.0(10)	92.0(10)
MWST-EM	1	29.0(10)	32.0(10)	32.0(10)	32.0(10)	32.0(10)
	2	27.0(10)	32.0(10)	32.0(10)	32.0(10)	32.0(10)
	3	28.0(10)	32.0(10)	32.0(10)	32.0(10)	32.0(10)
CC	1	<b>54.0(10)</b>	44.0(9)	36.0(9)	40.0(10)	60.0(10)
	2	32.0(5)	50.0(4)	50.0(5)	48.0(6)	40.0(10)
	3	5.0(4)	2.0(2)	39.0(2)	10.0(5)	54.0(7)
CC+T	1	14.0(10)	20.0(10)	18.0(10)	22.0(10)	26.0(10)
	2	20.0(7)	21.0(8)	19.0(10)	20.0(10)	23.0(10)
	3	17.0(4)	24.0(2)	19.0(2)	22.0(5)	16.0(9)
AC	1	40.0(10)	51.0(10)	53.0(10)	62.0(10)	73.0(10)
	2	36.0(10)	42.0(10)	52.0(10)	54.0(10)	66.0(10)
	3	30.0(10)	43.0(10)	46.0(10)	60.0(10)	66.0(10)
AC+T	1	24.0(10)	27.0(10)	29.0(10)	26.0(10)	26.0(10)
	2	22.0(10)	26.0(10)	28.0(10)	26.0(10)	24.0(10)
	3	24.0(10)	25.0(10)	26.0(10)	26.0(10)	26.0(10)
k10	1	50.0(10)	58.0(10)	72.0(10)	81.0(10)	89.0(10)
	2	<b>45.0(10)</b>	52.0(10)	70.0(10)	74.0(10)	<b>87.0(10)</b>
	3	<b>39.0(10)</b>	48.0(10)	<b>67.0(10)</b>	<b>70.0(10)</b>	<b>83.0(10)</b>
k10+T	1	23.0(10)	26.0(10)	28.0(10)	27.0(10)	22.0(10)
	2	20.0(10)	24.0(10)	28.0(10)	26.0(10)	22.0(10)
	3	20.0(10)	25.0(10)	26.0(10)	26.0(10)	25.0(10)
SEM	1	41.0(10)	45.0(10)	55.0(10)	68.0(10)	77.0(10)
	2	38.0(10)	38.0(10)	50.0(10)	60.0(10)	72.0(10)
	3	35.0(10)	38.0(10)	44.0(10)	53.0(10)	64.0(10)
SEM+T	1	50.0(10)†	<b>63.0(10)†</b>	<b>75.0(10)†</b>	<b>82.0(10)†</b>	<b>93.0(10)†</b>
	2	43.0(10)†	<b>57.0(10)†</b>	<b>71.0(10)†</b>	<b>76.0(10)†</b>	78.0(10)†
	3	38.0(10)†	<b>50.0(10)†</b>	60.0(10)†	66.0(10)†	75.0(10)†
SEM.bic	1	40.0(10)	49.0(10)	52.0(10)	65.0(10)	60.0(10)
	2	33.0(10)	32.0(10)	47.0(10)	60.0(10)	61.0(10)
	3	31.0(10)	36.0(10)	41.0(10)	49.0(10)	55.0(10)
SEM.bic+T	1	39.0(10)	50.0(10)	60.0(10)	69.0(10)	83.0(10)
	2	36.0(10)	46.0(10)	61.0(10)	67.0(10)	64.0(10)
	3	27.0(10)	39.0(10)	54.0(10)	53.0(10)	71.0(10)
GES-EM	1	43.0(8)	32.0(9)	46.0(10)	54.0(3)	-
	2	38.0(9)	41.0(10)	41.0(10)	-	-
	3	38.0(9)	40.0(9)	42.0(10)	42.0(10)	-
GES-EM+T	1	6.0(10)	9.0(10)	14.0(10)	18.0(9)	-
	2	9.0(10)	8.0(10)	15.0(10)	10.0(10)	-
	3	6.0(10)	9.0(10)	14.0(10)	14.0(9)	-
SEM+RT	1	40.0(10)	47.0(10)	54.0(10)	58.0(10)	63.0(10)
	2	27.0(10)	43.0(10)	44.0(10)	58.0(10)	60.0(10)
	3	30.0(10)	35.0(10)	45.0(10)	47.0(10)	56.0(10)

Recall Alarm1 nmar Recall in percentuale

Recall HailFinder nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	6.6(10)	7.2(10)	7.2(10)	8.4(10)	8.4(10)
MWST-EM	1	<b>12.0(10)</b>	<b>13.3(10)</b>	<b>12.7(10)</b>	<b>13.9(10)</b>	<b>22.3(10)</b>
	2	<b>9.6(10)</b>	<b>12.0(10)</b>	<b>12.7(10)</b>	<b>15.7(10)</b>	<b>15.7(10)</b>
	3	<b>10.8(10)</b>	<b>12.7(10)</b>	<b>13.3(10)</b>	<b>16.9(10)</b>	<b>16.9(10)</b>
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	6.0(10)	7.2(10)*	7.2(10)	8.4(10)	7.8(10)
	2	4.8(10)	7.2(10)	7.8(10)	7.2(10)	9.6(10)
	3	4.8(10)	6.0(10)	7.2(10)	7.2(10)	9.0(10)
k10+T	1	6.0(10)	6.6(10)	9.0(10)	10.2(10)	11.4(10)
	2	6.0(10)	6.6(10)	8.4(10)	10.8(10)	11.4(10)
	3	4.2(10)	6.0(10)	9.6(10)	10.2(10)	10.8(10)
SEM	1	3.6(10)	4.2(10)	7.8(10)	7.2(10)	10.8(7)
	2	2.4(10)	3.6(10)	6.6(10)	8.4(10)	9.6(5)
	3	2.4(10)	3.6(10)	5.4(10)	7.2(10)	8.4(7)
SEM+T	1	5.4(10)†	6.0(10)†	8.4(10)	10.2(10)†	12.0(7)*
	2	4.8(10)†	6.0(10)†	8.4(10)†	9.6(10)	12.0(5)
	3	4.2(10)†	6.0(10)†	8.4(10)†	8.4(10)*†	9.6(6)
SEM.bic	1	4.8(10)	4.8(10)	8.4(10)	8.4(10)	9.6(8)
	2	2.4(10)	4.2(10)	6.6(10)	7.2(10)	10.8(5)
	3	2.4(10)	3.6(10)	7.2(10)	6.6(10)	8.4(8)
SEM.bic+T	1	6.0(10)	6.6(10)	8.4(10)	8.4(10)	11.4(8)
	2	6.0(10)	6.0(10)	8.4(10)	9.0(10)	10.8(5)
	3	4.8(10)	6.6(10)	8.4(10)	9.0(10)	9.6(7)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	3.6(10)	4.2(10)	7.8(10)	8.4(10)	10.8(9)
	2	2.4(10)	4.8(10)	6.0(10)	7.8(10)	10.2(8)
	3	2.4(10)	3.6(10)	6.6(10)	7.2(10)	8.4(7)

Recall HailFinder nmar Recall in percentuale

Recall Hepar2 nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	10.8(10)	15.8(10)	26.2(10)	32.7(10)	48.1(10)
MWST-EM	1	<b>13.1(10)</b>	12.3(10)	13.1(10)	11.9(10)	13.1(10)
	2	<b>12.3(10)</b>	12.7(10)	13.1(10)	13.1(10)	13.1(10)
	3	<b>11.9(10)</b>	<b>12.7(10)</b>	13.1(10)	13.1(10)	13.1(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	9.2(10)	<b>13.1(10)</b>	<b>25.8(10)*</b>	<b>32.7(10)</b>	47.3(10)
	2	8.8(10)*	<b>13.1(10)*</b>	<b>25.0(10)*</b>	<b>29.2(10)</b>	46.2(9)
	3	8.5(10)*	11.5(10)	<b>21.5(10)*</b>	<b>26.9(10)</b>	<b>43.8(9)</b>
k10+T	1	5.0(10)	5.4(10)	7.7(10)	8.5(10)	10.4(10)
	2	4.6(10)	5.4(10)	7.7(10)	9.2(10)	10.8(10)
	3	4.2(10)	6.2(10)	5.8(10)	8.1(10)	9.6(10)
SEM	1	6.9(10)	10.8(10)	21.2(10)	30.4(10)	45.0(10)
	2	4.6(10)	8.8(10)	18.5(10)	27.7(10)	46.2(10)
	3	4.6(10)	8.5(10)	17.3(10)	25.0(10)	41.2(10)
SEM+T	1	7.3(10)	11.5(10)	21.2(10)	31.2(10)	44.2(10)
	2	5.4(10)	10.0(10)†	18.5(10)	29.2(10)	43.1(10)
	3	5.8(10)	9.6(10)†	17.7(10)†	25.8(10)	43.5(10)
SEM.bic	1	6.2(10)	10.8(10)	21.9(10)	28.5(10)	<b>48.1(10)</b>
	2	5.0(10)	9.6(10)	17.7(10)	27.7(10)	<b>46.5(10)</b>
	3	4.6(10)	7.7(10)	17.7(10)	23.5(10)	43.1(10)
SEM.bic+T	1	6.9(10)	11.2(10)	20.8(10)	30.0(10)	43.1(10)
	2	5.8(10)	10.0(10)	18.1(10)	28.1(10)	46.2(10)
	3	5.4(10)	8.8(10)	18.1(10)	23.5(10)	41.5(10)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	6.9(10)	10.8(10)	20.8(10)	30.8(10)	46.9(10)
	2	5.0(10)	8.8(10)	18.5(10)	27.3(10)	45.0(10)
	3	5.0(10)	9.2(10)	17.3(10)	24.6(10)	42.7(10)

Recall Hepar2 nmar Recall in percentuale

**Tempo**

		Tempo Asia mcar				
Metodo	p	100	200	500	1000	5000
MMHC	-	0.2(10)	0.2(10)	0.4(10)	0.9(10)	8.0(10)
	1	19.7(10)	36.7(10)	102.3(10)	207.0(10)	800.8(10)
MWST-EM	2	34.6(10)	52.1(10)	159.8(10)	288.4(10)	1507.2(10)
	3	32.7(10)	72.0(10)	170.1(10)	450.3(10)	2056.4(10)
	1	0.2(10)	0.2(10)	0.4(10)	0.7(10)	5.8(10)
CC	2	0.1(10)	0.2(10)	0.4(10)	0.7(10)	2.9(10)
	3	0.2(10)	0.2(10)	0.3(10)	0.6(10)	2.1(10)
	1	19.9(10)	37.0(10)	102.9(10)	207.7(10)	805.9(10)
CC+T	2	34.7(10)	52.3(10)	160.1(10)	288.9(10)	1509.9(10)
	3	32.8(10)	72.1(10)	170.4(10)	450.7(10)	2058.6(10)
	1	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
AC	2	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
	3	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
	1	19.8(10)	36.8(10)	102.4(10)	207.1(10)	800.9(10)
AC+T	2	34.6(10)	52.2(10)	159.8(10)	288.5(10)	1507.3(10)
	3	32.8(10)	72.0(10)	170.1(10)	450.4(10)	2056.5(10)
	1	0.1(10)*	0.2(10)*	0.2(10)*	0.4(10)*	2.9(10)*
k10	2	0.2(10)*	0.2(10)*	0.3(10)*	0.7(10)*	6.2(10)*
	3	0.2(10)*	0.3(10)*	0.5(10)*	1.1(10)*	7.9(10)*
	1	19.7(10)	36.8(10)	102.5(10)	207.3(10)	803.5(10)
k10+T	2	34.7(10)	52.3(10)	160.1(10)	288.9(10)	1513.5(10)
	3	32.8(10)	72.2(10)	170.5(10)	451.2(10)	2064.2(10)
	1	0.2(10)	0.3(10)	0.3(10)	0.4(10)	0.7(10)
SEM	2	0.2(10)	0.4(10)	0.6(10)	0.8(10)	1.9(10)
	3	0.4(10)	0.6(10)	0.7(10)	1.3(10)	3.6(10)
	1	19.9(10)	36.9(10)	102.6(10)	207.4(10)	801.4(10)
SEM+T	2	35.1(10)	52.6(10)	160.7(10)	289.3(10)	1508.8(10)
	3	33.3(10)	72.7(10)	170.9(10)	451.1(10)	2059.6(10)
	1	0.2(10)	0.2(10)	0.3(10)	0.4(10)	0.7(10)
SEM.bic	2	0.3(10)	0.4(10)	0.7(10)	0.9(10)	2.6(10)
	3	0.5(10)	0.8(10)	1.1(10)	2.4(10)	3.2(10)
	1	19.9(10)	37.0(10)	102.6(10)	207.3(10)	801.3(10)
SEM.bic+T	2	34.9(10)	52.5(10)	160.1(10)	289.2(10)	1509.3(10)
	3	33.2(10)	72.8(10)	171.1(10)	454.2(10)	2059.1(10)
	1	60.4(10)	164.8(10)	351.6(10)	671.0(10)	2974.2(10)
GES-EM	2	75.6(10)	124.5(10)	350.5(10)	713.3(10)	4217.8(10)
	3	66.4(10)	150.7(10)	386.3(10)	1076.1(10)	5454.1(10)
	1	96.0(10)	225.7(10)	511.3(10)	943.8(10)	4360.3(10)
GES-EM+T	2	106.3(10)	210.7(10)	648.3(10)	1146.5(10)	7334.8(10)
	3	135.8(10)	257.6(10)	722.6(10)	1602.4(10)	9587.9(10)
	1	0.3(10)†	0.2(10)†	0.2(10)†	0.3(10)†	0.7(10)†
SEM+RT	2	0.3(10)†	0.3(10)†	0.6(10)†	0.7(10)†	2.2(10)†
	3	0.4(10)†	0.4(10)†	0.6(10)†	1.3(10)†	3.9(10)†

**Tempo Asia mcar** Tempo in secondi

Tempo Child mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	0.7(10)	0.6(10)	1.3(10)	2.7(10)	38.0(10)
MWST-EM	1	103.5(10)	235.4(10)	545.1(10)	736.5(10)	3434.1(10)
	2	148.9(10)	222.4(10)	505.0(10)	1002.5(10)	3565.2(10)
	3	132.0(10)	209.1(10)	668.6(10)	1001.1(10)	3660.5(10)
CC	1	1.1(10)	0.6(10)	0.8(10)	1.4(10)	12.5(10)
	2	4.7(10)	6.9(10)	1.6(10)	1.2(10)	4.0(10)
	3	4.0(3)	4.4(4)	6.0(7)	33.8(10)	<b>4.0(10)</b>
CC+T	1	103.8(10)	235.7(10)	545.7(10)	737.5(10)	3447.7(10)
	2	149.2(10)	222.6(10)	505.4(10)	1003.2(10)	3568.9(10)
	3	138.3(3)	215.9(4)	695.6(7)	1001.7(10)	3663.3(10)
AC	1	0.5(10)	<b>0.4(10)</b>	<b>0.5(10)</b>	<b>0.8(10)</b>	<b>3.8(10)</b>
	2	1.7(10)	1.6(10)	<b>0.6(10)</b>	<b>0.8(10)</b>	<b>3.6(10)</b>
	3	1.8(10)	0.9(10)	<b>0.8(10)</b>	<b>0.9(10)</b>	4.0(10)
AC+T	1	103.7(10)	235.5(10)	545.3(10)	736.6(10)	3434.3(10)
	2	149.0(10)	222.5(10)	505.1(10)	1002.7(10)	3565.5(10)
	3	132.1(10)	209.3(10)	668.7(10)	1001.2(10)	3660.7(10)
k10	1	<b>0.4(10)*</b>	0.5(10)*	0.8(10)*	1.3(10)*	12.3(10)*
	2	<b>0.5(10)*</b>	<b>0.5(10)*</b>	1.0(10)*	2.1(10)*	22.0(10)*
	3	<b>0.6(10)*</b>	<b>0.6(10)*</b>	1.4(10)*	3.0(10)*	34.3(10)*
k10+T	1	103.7(10)	235.6(10)	545.5(10)	737.0(10)	3441.6(10)
	2	149.1(10)	222.6(10)	505.5(10)	1003.6(10)	3577.1(10)
	3	132.2(10)	209.5(10)	669.3(10)	1002.6(10)	3674.3(10)
SEM	1	1.4(10)	1.8(10)	7.3(10)	11.3(10)	66.9(10)
	2	2.5(10)	6.7(10)	11.2(10)	28.5(10)	144.0(10)
	3	2.1(10)	7.9(10)	23.2(10)	41.1(10)	251.3(10)
SEM+T	1	105.3(10)	237.9(10)	550.8(10)	746.7(10)	3505.0(10)
	2	152.3(10)	227.6(10)	514.4(10)	1033.1(10)	3689.6(10)
	3	137.2(10)	220.9(10)	686.5(10)	1051.8(10)	4072.2(9)
SEM.bic	1	1.4(10)	1.9(10)	5.9(10)	10.9(10)	69.1(10)
	2	2.5(10)	5.4(10)	12.6(10)	29.1(10)	153.7(10)
	3	1.9(10)	7.3(10)	22.2(10)	28.8(10)	251.6(10)
SEM.bic+T	1	104.7(10)	237.3(10)	552.9(10)	748.2(10)	3505.6(10)
	2	152.1(10)	228.7(10)	519.4(10)	1031.5(10)	3717.7(9)
	3	135.4(10)	218.9(10)	690.9(10)	1056.4(10)	3914.1(9)
GES-EM	1	601.0(10)	1198.8(10)	3893.0(10)	8824.9(10)	37481.3(10)
	2	834.1(10)	1442.8(10)	3383.2(10)	8590.9(10)	34559.4(9)
	3	256.6(10)	1292.9(10)	4376.3(10)	6673.0(10)	38087.0(10)
GES-EM+T	1	591.5(10)	1337.3(9)	3537.6(10)	7450.3(10)	41727.9(10)
	2	567.0(10)	1033.1(10)	3485.0(10)	7494.9(6)	43531.8(8)
	3	-	-	-	-	43845.6(3)
SEM+RT	1	1.5(10)†	1.9(10)†	7.5(10)†	11.7(10)†	77.8(10)†
	2	2.4(10)†	6.4(10)†	12.8(10)†	24.6(10)†	160.2(10)†
	3	2.5(10)†	7.7(10)†	18.4(10)†	53.3(10)†	232.9(10)†

Tempo Child mcar Tempo in secondi

Tempo Insurance mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	1.6(10)	1.4(10)	2.2(10)	4.1(10)	53.7(10)
MWST-EM	1	157.2(10)	272.0(10)	571.1(10)	1124.8(10)	6135.9(10)
	2	224.3(10)	403.3(10)	730.8(10)	1820.4(10)	6844.4(10)
	3	192.1(10)	307.8(10)	726.0(10)	1420.8(10)	6761.9(10)
CC	1	3.1(10)	2.0(10)	1.6(10)	2.0(10)	15.0(10)
	2	7657.9(5)	2569.0(7)	161.1(10)	44.2(10)	5.3(10)
	3	-	5816.5(3)	8451.7(1)	9191.8(1)	2665.2(9)
CC+T	1	157.8(10)	272.7(10)	572.4(10)	1126.8(10)	6148.8(10)
	2	224.8(6)	403.7(10)	731.6(10)	1821.5(10)	6848.3(10)
	3	-	305.9(3)	679.4(1)	1967.8(1)	6765.2(10)
AC	1	1.5(10)	<b>1.1(10)</b>	<b>1.0(10)</b>	<b>1.4(10)</b>	<b>4.3(10)</b>
	2	426.9(10)	2.9(10)	<b>1.2(10)</b>	<b>1.3(10)</b>	<b>4.0(10)</b>
	3	102.2(10)	5.8(10)	<b>1.3(10)</b>	<b>1.2(10)</b>	<b>3.0(10)</b>
AC+T	1	157.5(10)	272.3(10)	571.4(10)	1125.2(10)	6136.3(10)
	2	224.6(10)	403.6(10)	731.1(10)	1820.7(10)	6844.7(10)
	3	192.4(10)	308.0(10)	726.3(10)	1421.2(10)	6762.2(10)
k10	1	<b>1.2(10)*</b>	1.3(10)*	1.6(10)*	2.7(10)*	21.0(10)*
	2	<b>1.4(10)*</b>	<b>1.5(10)*</b>	2.1(10)*	4.4(10)*	43.7(10)*
	3	<b>1.5(10)*</b>	<b>1.5(10)*</b>	2.7(10)*	6.0(10)*	81.4(10)
k10+T	1	157.6(10)	272.6(10)	571.7(10)	1125.9(10)	6145.6(10)
	2	224.8(10)	403.9(10)	731.7(10)	1822.4(10)	6859.0(10)
	3	192.5(10)	308.5(10)	727.5(10)	1423.7(10)	6779.3(10)
SEM	1	2.3(10)	5.0(10)	14.0(10)	28.5(10)	180.8(10)
	2	6.0(10)	14.3(10)	33.1(10)	88.2(10)	495.9(9)
	3	6.4(10)	16.6(10)	42.9(10)	102.6(10)	1049.4(4)
SEM+T	1	160.7(10)	277.8(10)	584.3(10)	1154.4(10)	6406.3(10)
	2	234.4(10)	414.0(10)	769.6(10)	1937.5(10)	7780.6(7)
	3	202.2(10)	323.7(10)	791.3(10)	1602.8(10)	4730.7(1)
SEM.bic	1	2.2(10)	4.1(10)	10.4(10)	26.1(10)	200.3(10)
	2	5.4(10)	11.9(10)	26.0(10)	55.5(10)	421.2(9)
	3	6.7(10)	19.1(10)	43.7(10)	96.7(10)	607.3(6)
SEM.bic+T	1	160.0(10)	276.3(10)	583.1(10)	1151.7(10)	6291.3(10)
	2	229.6(10)	414.6(10)	760.2(10)	1917.6(10)	7449.4(8)
	3	199.9(10)	323.4(10)	794.0(10)	1595.8(10)	9788.7(2)
GES-EM	1	1650.8(10)	3393.2(10)	9785.8(9)	26570.6(9)	-
	2	1784.6(10)	3583.9(10)	24534.8(9)	22703.2(9)	-
	3	2956.1(9)	7688.3(10)	17099.2(10)	27439.3(10)	-
GES-EM+T	1	1464.6(10)	3107.9(10)	7683.5(10)	16907.8(10)	-
	2	1383.0(10)	4102.7(10)	11024.3(10)	30436.3(10)	-
	3	1126.3(10)	4671.1(10)	18560.3(10)	39432.7(8)	-
SEM+RT	1	2.7(10)†	5.5(10)†	13.9(10)†	34.9(10)†	187.3(10)†
	2	4.9(10)†	17.2(10)†	28.2(10)†	86.1(10)†	687.1(7)
	3	10.2(10)†	24.4(10)†	52.9(10)†	190.5(9)†	750.5(2)

Tempo Insurance mcar Tempo in secondi

Tempo Alarm1 mcar						
Metodo	p	100	200	500	1000	5000
MMHC	-	4.4(10)	4.1(10)	6.0(10)	8.0(10)	60.0(10)
MWST-EM	1	195.4(10)	278.3(10)	801.4(10)	814.5(10)	3846.3(10)
	2	211.1(10)	387.7(10)	941.2(10)	1622.6(10)	4614.7(10)
	3	245.9(10)	457.3(10)	736.0(10)	1884.6(10)	5823.2(10)
CC	1	372.4(10)	97.3(10)	5.1(10)	5.6(10)	14.1(10)
	2	20.1(3)	22.5(4)	12144.1(2)	29.5(5)	239.7(10)
	3	-	-	6.3(1)	17.5(1)	22.3(1)
CC+T	1	196.1(10)	279.0(10)	802.4(10)	816.2(10)	3855.8(10)
	2	102.2(4)	388.4(6)	948.2(3)	1623.8(10)	4619.1(10)
	3	-	-	2058.9(1)	934.5(1)	11860.8(1)
AC	1	<b>4.5(10)</b>	<b>3.3(10)</b>	<b>3.4(10)</b>	<b>4.2(10)</b>	<b>8.7(10)</b>
	2	648.4(10)	6.4(10)	<b>6.6(10)</b>	<b>4.7(10)</b>	<b>7.1(10)</b>
	3	460.4(10)	6.6(10)	<b>4.0(10)</b>	<b>3.7(10)</b>	<b>5.5(10)</b>
AC+T	1	195.8(10)	278.8(10)	801.8(10)	815.0(10)	3846.9(10)
	2	211.5(10)	388.1(10)	941.5(10)	1623.0(10)	4615.3(10)
	3	246.2(10)	457.6(10)	736.4(10)	1885.0(10)	5823.8(10)
k10	1	6.9(10)*	5.5(10)*	5.1(10)*	6.4(10)*	25.0(10)*
	2	<b>4.4(10)*</b>	<b>4.3(10)*</b>	7.6(10)*	10.7(10)*	44.1(10)*
	3	<b>5.2(10)*</b>	<b>5.8(10)*</b>	10.6(10)*	17.7(10)*	77.9(10)
k10+T	1	196.0(10)	278.8(10)	802.2(10)	815.8(10)	3860.1(10)
	2	211.8(10)	388.4(10)	942.5(10)	1624.7(10)	4633.0(10)
	3	246.6(10)	458.3(10)	737.7(10)	1887.8(10)	5845.0(10)
SEM	1	7.2(10)	13.8(10)	32.9(10)	76.8(10)	424.6(10)
	2	12.0(10)	25.1(10)	69.2(10)	219.6(9)	1204.9(8)
	3	24.8(10)	43.2(10)	137.7(10)	232.4(8)	1069.7(1)
SEM+T	1	201.8(10)	299.7(10)	826.3(10)	878.0(10)	4161.1(10)
	2	220.7(10)	434.1(10)	1057.8(10)	1868.7(10)	5840.7(8)
	3	264.4(10)	524.1(10)	870.7(10)	2226.2(9)	9738.3(5)
SEM.bic	1	5.6(10)	10.0(10)	27.0(10)	55.0(10)	333.4(10)
	2	10.5(10)	21.3(10)	83.2(10)	193.1(10)	1037.5(7)
	3	18.0(10)	41.6(10)	130.0(10)	310.7(10)	2750.4(3)
SEM.bic+T	1	202.7(10)	289.2(10)	833.7(10)	857.4(10)	4114.0(10)
	2	221.5(10)	407.8(10)	1019.4(10)	1835.1(10)	5394.5(10)
	3	264.3(10)	507.2(10)	866.2(10)	2183.6(9)	8720.4(6)
GES-EM	1	7528.3(10)	11387.8(9)	26904.6(10)	-	-
	2	11311.4(10)	11425.6(10)	28636.2(10)	41117.1(6)	-
	3	8802.9(10)	11261.5(10)	22247.5(10)	37163.9(10)	-
GES-EM+T	1	2604.1(10)	7410.3(10)	23558.8(10)	38566.1(7)	-
	2	3687.3(10)	13777.7(10)	-	43318.4(3)	-
	3	7524.2(10)	14907.0(10)	39791.8(9)	-	-
SEM+RT	1	6.1(10)†	13.9(10)†	35.8(10)†	71.6(10)†	478.4(9)†
	2	12.9(10)†	30.9(10)†	104.5(10)†	229.9(9)†	1069.7(5)
	3	24.2(10)†	46.2(10)†	173.5(10)†	195.2(7)†	-

Tempo Alarm1 mcar Tempo in secondi

Tempo HailFinder mcar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	10.4(10)	5.8(10)	6.9(10)	11.5(10)	92.2(10)
MWST-EM	1	364.1(10)	678.1(10)	1183.5(10)	2422.7(10)	7336.8(10)
	2	400.7(10)	664.3(10)	1489.0(10)	2889.9(10)	13419.1(10)
	3	447.9(10)	774.3(10)	1440.5(10)	2759.5(10)	13462.2(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	<b>7.7(10)*</b>	<b>4.6(10)*</b>	<b>5.1(10)*</b>	<b>6.6(10)*</b>	<b>39.4(10)*</b>
	2	<b>8.2(10)*</b>	<b>5.4(10)*</b>	<b>6.5(10)*</b>	<b>10.1(10)*</b>	<b>62.6(10)</b>
	3	<b>9.1(10)*</b>	<b>6.6(10)*</b>	<b>8.5(10)*</b>	<b>16.0(10)*</b>	<b>136.4(10)</b>
k10+T	1	365.6(10)	679.6(10)	1185.2(10)	2425.2(10)	7358.0(10)
	2	402.3(10)	666.0(10)	1491.5(10)	2893.6(10)	13445.4(10)
	3	449.6(10)	776.4(10)	1443.5(10)	2764.9(10)	13494.3(10)
SEM	1	15.3(10)	33.9(10)	109.9(10)	196.0(10)	889.9(8)
	2	37.9(10)	86.8(10)	324.3(10)	572.1(8)	5654.4(1)
	3	51.0(10)	169.0(10)	565.6(10)	957.2(9)	-
SEM+T	1	377.1(10)	722.4(10)	1326.3(10)	2649.7(10)	9431.7(8)
	2	471.3(10)	792.0(10)	1763.4(10)	3421.7(8)	16283.3(2)
	3	540.7(10)	984.7(10)	2082.8(10)	4114.4(10)	-
SEM.bic	1	14.6(10)	37.2(10)	95.0(10)	174.8(10)	947.5(5)
	2	34.6(10)	81.5(10)	275.9(10)	676.6(6)	-
	3	54.6(10)	150.3(10)	458.9(10)	1041.5(7)	-
SEM.bic+T	1	377.2(10)	712.8(10)	1306.9(10)	2599.3(10)	8465.0(6)
	2	439.5(10)	793.6(10)	1779.4(10)	3571.1(7)	-
	3	525.2(10)	988.9(10)	2059.3(10)	4204.3(9)	-
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	19.3(10)†	44.1(10)†	185.5(10)†	187.0(10)†	1236.4(8)†
	2	42.2(10)†	89.5(10)†	365.5(10)†	860.9(8)†	-
	3	65.9(10)†	202.2(10)†	577.1(9)†	1366.5(8)†	-

Tempo HailFinder mcar Tempo in secondi

		Tempo Hepar2 mcar				
Metodo	p	100	200	500	1000	5000
MMHC	-	22.4(10)	38.9(10)	73.6(10)	382.9(10)	531.6(10)
MWST-EM	1	403.4(10)	685.0(10)	1734.6(10)	2796.6(10)	6708.3(10)
	2	493.0(10)	777.8(10)	1926.4(10)	3075.7(10)	12047.1(10)
	3	442.4(10)	790.9(10)	1431.1(10)	2340.5(10)	8950.3(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	20.9(10)*	<b>37.8(10)*</b>	<b>92.2(10)*</b>	551.6(10)*	<b>667.6(10)*</b>
	2	<b>27.5(10)*</b>	<b>50.2(10)*</b>	<b>164.9(10)*</b>	<b>728.5(10)*</b>	4842.6(10)*
	3	<b>29.7(10)*</b>	<b>73.0(10)*</b>	<b>361.2(10)*</b>	1519.3(10)*	<b>4342.6(10)</b>
k10+T	1	406.1(10)	687.8(10)	1737.9(10)	2801.2(10)	6734.0(10)
	2	496.2(10)	781.1(10)	1930.4(10)	3081.6(10)	12079.4(10)
	3	445.7(10)	794.7(10)	1435.9(10)	2348.2(10)	8989.3(10)
SEM	1	18.1(10)	38.6(10)	127.6(10)	266.9(10)	1681.9(10)
	2	46.4(10)	115.0(10)	596.3(10)	753.6(10)	5454.5(10)
	3	31.9(10)	166.9(10)	482.6(10)	1229.0(10)	8279.0(10)
SEM+T	1	424.8(10)	737.4(10)	1855.0(10)	3015.7(10)	8551.0(10)
	2	561.4(10)	920.5(10)	2456.9(10)	4088.9(10)	17619.4(10)
	3	619.5(10)	995.3(10)	2147.8(10)	4165.4(10)	19459.5(10)
SEM.bic	1	<b>18.0(10)</b>	39.9(10)	128.5(10)	<b>244.5(10)</b>	1589.2(10)
	2	48.3(10)	109.9(10)	332.5(10)	733.1(10)	<b>3769.9(10)</b>
	3	38.3(10)	141.0(10)	453.0(10)	<b>1116.7(10)</b>	6215.2(10)
SEM.bic+T	1	424.6(10)	731.3(10)	1867.4(10)	3033.1(10)	8177.3(10)
	2	542.7(10)	889.5(10)	2354.6(10)	3757.7(10)	16685.2(10)
	3	551.9(10)	982.9(10)	2157.2(10)	3665.4(10)	19628.1(10)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	19.5(10)†	44.3(10)†	121.7(10)†	271.2(10)†	1353.8(10)†
	2	54.4(10)†	119.9(10)†	670.5(10)†	796.0(10)†	6055.3(10)†
	3	51.4(10)†	168.7(10)†	511.7(10)†	1155.2(10)†	7821.6(10)†

Tempo Hepar2 mcar    Tempo in secondi

		Tempo Asia mar				
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	<i>0.2(10)</i>	<i>0.2(10)</i>	<i>0.4(10)</i>	<i>0.9(10)</i>	<i>8.0(10)</i>
MWST-EM	1	30.2(10)	42.3(10)	139.0(10)	248.8(10)	1027.1(10)
	2	31.8(10)	58.6(10)	180.3(10)	283.0(10)	1360.5(10)
	3	31.8(10)	72.1(10)	201.6(10)	326.0(10)	1552.7(10)
CC	1	0.2(10)	0.2(10)	0.4(10)	0.7(10)	7.0(10)
	2	0.1(10)	0.2(10)	0.4(10)	0.7(10)	4.7(10)
	3	0.2(10)	0.2(10)	0.4(10)	0.6(10)	3.1(10)
CC+T	1	30.4(10)	42.5(10)	139.5(10)	249.7(10)	1032.9(10)
	2	32.0(10)	58.8(10)	180.6(10)	283.8(10)	1365.1(10)
	3	31.9(10)	72.4(10)	201.9(10)	326.9(10)	1556.3(10)
AC	1	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
	2	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
	3	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
AC+T	1	30.3(10)	42.4(10)	139.1(10)	248.9(10)	1027.2(10)
	2	31.9(10)	58.6(10)	180.3(10)	283.1(10)	1360.6(10)
	3	31.8(10)	72.2(10)	201.7(10)	326.1(10)	1552.8(10)
k10	1	0.1(10)*	0.1(10)*	0.3(10)*	0.4(10)*	1.4(10)*
	2	0.1(10)*	0.2(10)*	0.3(10)*	0.5(10)*	3.8(10)*
	3	0.2(10)*	0.2(10)*	0.3(10)*	0.7(10)*	5.9(10)*
k10+T	1	30.3(10)	42.4(10)	139.2(10)	249.2(10)	1028.3(10)
	2	31.9(10)	58.7(10)	180.6(10)	283.5(10)	1364.1(10)
	3	31.9(10)	72.3(10)	201.9(10)	326.5(10)	1557.3(10)
SEM	1	0.2(10)	0.2(10)	0.2(10)	0.3(10)	0.3(10)
	2	0.2(10)	0.2(10)	0.2(10)	0.3(10)	0.4(10)
	3	0.1(10)	0.2(10)	0.2(10)	0.3(10)	0.4(10)
SEM+T	1	30.4(10)	42.5(10)	139.2(10)	249.1(10)	1027.3(10)
	2	32.0(10)	58.8(10)	180.5(10)	283.4(10)	1360.9(10)
	3	31.9(10)	72.4(10)	201.8(10)	326.4(10)	1553.1(10)
SEM.bic	1	0.2(10)	0.2(10)	0.2(10)	0.3(10)	0.3(10)
	2	0.2(10)	0.2(10)	0.2(10)	0.3(10)	0.4(10)
	3	0.2(10)	0.2(10)	0.2(10)	0.2(10)	0.5(10)
SEM.bic+T	1	30.4(10)	42.5(10)	139.2(10)	249.1(10)	1027.4(10)
	2	32.0(10)	58.7(10)	180.5(10)	283.3(10)	1361.0(10)
	3	31.9(10)	72.3(10)	201.8(10)	326.3(10)	1553.1(10)
GES-EM	1	73.7(10)	152.9(10)	327.7(10)	645.2(10)	2885.2(10)
	2	54.4(10)	114.3(10)	347.3(9)	685.7(10)	3287.7(9)
	3	56.0(10)	106.6(10)	343.5(10)	759.9(10)	3720.6(10)
GES-EM+T	1	88.9(9)	208.2(10)	444.7(10)	1070.3(10)	3652.0(10)
	2	116.3(10)	215.9(10)	555.0(10)	1070.6(10)	4993.4(9)
	3	105.3(9)	248.1(10)	678.7(10)	1343.8(10)	4711.6(10)
SEM+RT	1	<i>0.4(10)†</i>	<i>0.2(10)†</i>	<i>0.2(10)†</i>	<i>0.3(10)†</i>	<i>0.3(10)†</i>
	2	<i>0.2(10)†</i>	<i>0.2(10)†</i>	<i>0.2(10)†</i>	<i>0.3(10)†</i>	<i>0.3(10)†</i>
	3	<i>0.1(10)†</i>	<i>0.2(10)†</i>	<i>0.2(10)†</i>	<i>0.3(10)†</i>	<i>0.4(10)†</i>

**Tempo Asia mar** Tempo in secondi

Tempo Child mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	0.7(10)	0.6(10)	1.3(10)	2.7(10)	38.0(10)
MWST-EM	1	135.2(10)	249.4(10)	695.4(10)	1120.1(10)	6886.6(10)
	2	135.8(10)	233.3(10)	544.2(10)	1038.1(10)	5196.1(10)
	3	113.8(10)	252.3(10)	556.8(10)	995.8(10)	3979.1(10)
CC	1	1.7(10)	0.8(10)	0.7(10)	1.3(10)	9.7(10)
	2	19.1(10)	1.5(10)	0.9(10)	1.0(10)	5.0(10)
	3	10.7(9)	3.9(10)	2.5(10)	1.8(10)	3.9(10)
CC+T	1	135.4(10)	249.7(10)	695.8(10)	1121.0(10)	6893.5(10)
	2	136.0(10)	233.6(10)	544.6(10)	1038.8(10)	5200.5(10)
	3	119.5(9)	252.6(10)	557.2(10)	996.5(10)	3982.5(10)
AC	1	0.8(10)	<b>0.5(10)</b>	<b>0.5(10)</b>	<b>0.8(10)</b>	<b>3.6(10)</b>
	2	0.8(10)	0.6(10)	<b>0.6(10)</b>	<b>0.7(10)</b>	<b>3.8(10)</b>
	3	0.8(10)	0.6(10)	<b>0.6(10)</b>	<b>0.8(10)</b>	<b>3.6(10)</b>
AC+T	1	135.3(10)	249.6(10)	695.5(10)	1120.2(10)	6886.9(10)
	2	135.9(10)	233.5(10)	544.3(10)	1038.2(10)	5196.4(10)
	3	114.0(10)	252.4(10)	556.9(10)	996.0(10)	3979.3(10)
k10	1	<b>0.4(10)*</b>	0.5(10)*	0.8(10)*	1.5(10)*	17.5(10)*
	2	<b>0.5(10)*</b>	<b>0.5(10)*</b>	1.0(10)*	2.0(10)*	21.7(10)*
	3	<b>0.6(10)*</b>	<b>0.6(10)*</b>	1.2(10)*	2.4(10)*	27.8(10)*
k10+T	1	135.4(10)	249.7(10)	695.7(10)	1120.8(10)	6896.5(10)
	2	136.0(10)	233.6(10)	544.7(10)	1039.1(10)	5207.4(10)
	3	114.1(10)	252.6(10)	557.4(10)	997.0(10)	3991.9(10)
SEM	1	1.0(10)	2.0(10)	5.6(10)	9.0(10)	65.5(10)
	2	1.8(10)	2.8(10)	6.9(10)	18.5(10)	118.1(10)
	3	1.6(10)	3.7(10)	8.5(10)	19.9(10)	181.7(10)
SEM+T	1	137.1(10)	251.7(10)	700.5(10)	1128.0(10)	6957.0(10)
	2	137.6(10)	236.2(10)	550.4(10)	1074.0(10)	5300.6(10)
	3	115.7(10)	255.1(10)	566.8(10)	1030.2(10)	4123.9(10)
SEM.bic	1	1.0(10)	2.1(10)	6.3(10)	11.0(10)	94.5(10)
	2	1.3(10)	2.7(10)	7.0(10)	17.7(10)	111.8(10)
	3	1.4(10)	1.9(10)	7.3(10)	13.7(10)	137.5(10)
SEM.bic+T	1	138.6(10)	252.0(10)	703.3(10)	1134.3(10)	6975.6(10)
	2	137.9(10)	235.1(10)	555.1(10)	1069.2(10)	5341.0(10)
	3	116.9(10)	254.8(10)	568.9(10)	1016.6(10)	4126.6(10)
GES-EM	1	531.6(10)	1196.1(10)	3866.9(10)	9553.4(10)	38167.7(8)
	2	749.2(10)	1289.4(10)	4227.7(10)	6411.0(10)	37347.8(10)
	3	426.2(10)	1062.4(10)	2805.2(10)	6120.5(10)	37464.2(9)
GES-EM+T	1	627.5(10)	1251.6(8)	3623.2(9)	7179.3(10)	46851.9(9)
	2	542.5(10)	1278.5(10)	3625.9(10)	7313.8(7)	45682.7(6)
	3	-	-	-	-	38884.2(7)
SEM+RT	1	1.0(10)†	2.0(10)†	4.2(10)†	12.3(10)†	80.0(10)†
	2	1.6(10)†	2.5(10)†	8.4(10)†	21.1(10)†	106.7(10)†
	3	1.7(10)†	3.5(10)†	7.7(10)†	19.0(10)†	94.4(10)†

**Tempo Child mar** Tempo in secondi

Metodo	p	Tempo Insurance mar				
		100	200	500	1000	5000
MMHC	-	1.6(10)	1.4(10)	2.2(10)	4.1(10)	53.7(10)
MWST-EM	1	156.4(10)	282.7(10)	858.2(10)	1388.5(10)	9337.3(10)
	2	267.9(10)	353.0(10)	798.8(10)	1894.9(10)	7020.7(10)
	3	167.5(10)	304.4(10)	915.2(10)	1404.4(10)	6680.8(10)
CC	1	8.0(10)	3.4(10)	1.9(10)	2.1(10)	9.9(10)
	2	52.0(4)	11.3(9)	13.3(10)	2.6(9)	6.5(10)
	3	3700.4(4)	1155.3(8)	2019.6(8)	445.6(6)	5.3(10)
CC+T	1	157.0(10)	283.2(10)	859.0(10)	1390.2(10)	9347.0(10)
	2	263.8(9)	353.6(10)	799.6(10)	1895.9(10)	7025.5(10)
	3	157.2(5)	306.2(9)	916.3(10)	1405.6(10)	6684.7(10)
AC	1	1.5(10)	1.3(10)	<b>1.0(10)</b>	<b>1.3(10)</b>	<b>4.0(10)</b>
	2	4.9(10)	1.7(10)	<b>1.1(10)</b>	<b>1.6(10)</b>	<b>4.2(10)</b>
	3	75.0(10)	1.7(10)	<b>1.1(10)</b>	<b>1.2(10)</b>	<b>3.8(10)</b>
AC+T	1	156.9(10)	283.0(10)	858.5(10)	1388.9(10)	9337.6(10)
	2	268.3(10)	353.3(10)	799.2(10)	1895.3(10)	7021.1(10)
	3	167.9(10)	304.7(10)	915.5(10)	1404.7(10)	6681.2(10)
k10	1	<b>1.2(10)*</b>	<b>1.2(10)*</b>	1.6(10)*	3.0(10)*	24.7(10)*
	2	<b>1.3(10)*</b>	<b>1.3(10)*</b>	1.9(10)*	3.6(10)*	34.4(10)*
	3	<b>1.4(10)*</b>	<b>1.3(10)*</b>	2.2(10)*	5.4(10)*	47.2(10)*
k10+T	1	156.9(10)	283.2(10)	859.0(10)	1389.6(10)	9348.6(10)
	2	268.4(10)	353.7(10)	799.7(10)	1896.8(10)	7034.6(10)
	3	168.0(10)	305.2(10)	916.4(10)	1407.0(10)	6696.2(10)
SEM	1	2.3(10)	4.9(10)	10.7(10)	30.2(10)	226.9(10)
	2	3.2(10)	6.4(10)	12.7(10)	48.7(10)	384.1(10)
	3	3.3(10)	7.3(10)	21.2(10)	40.2(10)	400.2(10)
SEM+T	1	161.5(10)	288.1(10)	869.1(10)	1448.9(10)	9541.5(10)
	2	270.8(10)	359.8(10)	818.6(10)	1940.2(10)	7467.8(10)
	3	172.4(10)	312.3(10)	942.1(10)	1498.5(10)	7110.5(10)
SEM.bic	1	2.3(10)	4.0(10)	10.8(10)	22.7(10)	183.1(10)
	2	3.4(10)	5.5(10)	15.0(10)	39.4(10)	285.2(10)
	3	3.4(10)	7.7(10)	16.8(10)	45.0(10)	302.6(10)
SEM.bic+T	1	159.9(10)	286.9(10)	874.7(10)	1420.2(10)	9526.6(10)
	2	273.0(10)	359.0(10)	821.0(10)	1929.3(10)	7453.9(10)
	3	170.4(10)	310.2(10)	948.6(10)	1447.9(10)	7110.0(10)
GES-EM	1	1594.4(10)	3143.5(10)	14691.7(10)	28379.4(10)	-
	2	1460.4(10)	2955.1(10)	20539.9(9)	19431.2(7)	-
	3	1287.4(10)	4326.2(10)	11314.0(10)	19717.3(8)	-
GES-EM+T	1	1597.4(9)	2313.8(9)	7023.4(10)	14146.4(10)	-
	2	1185.4(9)	2201.8(10)	6295.3(10)	13966.3(10)	-
	3	1032.8(9)	2184.5(10)	6460.5(10)	16325.0(9)	-
SEM+RT	1	2.6(10)†	5.0(10)†	12.4(10)†	31.1(10)†	279.7(10)†
	2	3.7(10)†	6.8(10)†	16.3(10)†	45.2(10)†	391.2(10)†
	3	3.7(10)†	8.3(10)†	17.9(10)†	57.3(10)†	352.1(10)†

Tempo Insurance mar Tempo in secondi

Tempo Alarm1 mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	4.4(10)	4.1(10)	6.0(10)	8.0(10)	60.0(10)
MWST-EM	1	211.6(10)	408.5(10)	697.6(10)	991.8(10)	4678.9(10)
	2	207.5(10)	197.5(10)	740.6(10)	1460.7(10)	4564.8(10)
	3	234.8(10)	372.7(10)	522.3(10)	996.2(10)	4996.4(10)
CC	1	8032.8(6)	1709.2(10)	7.4(10)	9.9(10)	10.0(10)
	2	18.9(4)	7792.6(2)	5562.8(10)	403.0(6)	9.7(10)
	3	17.6(2)	27.2(3)	11095.9(1)	1557.5(6)	503.6(7)
CC+T	1	212.1(10)	409.3(10)	698.5(10)	993.1(10)	4687.0(10)
	2	206.7(5)	203.5(7)	741.5(10)	1462.0(10)	4569.4(10)
	3	101.4(3)	341.4(4)	476.1(1)	988.7(7)	4806.2(8)
AC	1	131.4(10)	<b>3.3(10)</b>	<b>3.5(10)</b>	<b>3.6(10)</b>	<b>8.3(10)</b>
	2	543.7(10)	<b>4.0(10)</b>	<b>4.0(10)</b>	<b>4.1(10)</b>	<b>8.4(10)</b>
	3	697.5(10)	8.2(10)	<b>3.7(10)</b>	<b>3.6(10)</b>	<b>6.8(10)</b>
AC+T	1	212.0(10)	409.0(10)	698.0(10)	992.2(10)	4679.5(10)
	2	207.9(10)	197.9(10)	741.0(10)	1461.1(10)	4565.4(10)
	3	235.2(10)	373.1(10)	522.8(10)	996.6(10)	4997.0(10)
k10	1	5.7(10)*	5.9(10)*	5.2(10)*	7.8(10)*	32.3(10)*
	2	<b>4.4(10)*</b>	4.6(10)*	7.0(10)*	8.1(10)*	39.3(10)*
	3	<b>4.7(10)*</b>	<b>4.5(10)*</b>	7.3(10)*	10.2(10)*	56.2(10)*
k10+T	1	212.1(10)	409.1(10)	698.5(10)	993.5(10)	4693.5(10)
	2	208.1(10)	198.4(10)	742.1(10)	1462.7(10)	4582.5(10)
	3	235.4(10)	373.5(10)	523.5(10)	998.7(10)	5016.9(10)
SEM	1	4.8(10)	11.3(10)	25.9(10)	79.8(10)	430.0(10)
	2	5.3(10)	13.0(10)	40.5(10)	79.6(10)	606.6(10)
	3	7.3(10)	13.8(10)	44.1(10)	108.6(10)	811.4(10)
SEM+T	1	217.8(10)	420.1(10)	736.6(10)	1062.7(10)	5374.6(10)
	2	216.5(10)	221.3(9)	811.6(10)	1551.7(10)	5427.1(10)
	3	242.7(10)	360.9(9)	615.9(10)	1095.9(10)	5843.2(10)
SEM.bic	1	<b>4.3(10)</b>	8.6(10)	20.9(10)	71.2(10)	346.4(10)
	2	6.5(10)	13.4(10)	32.3(10)	75.4(10)	474.5(10)
	3	6.3(10)	11.3(10)	48.5(10)	89.7(10)	594.4(10)
SEM.bic+T	1	216.3(10)	420.2(10)	725.4(10)	1084.8(10)	4975.4(10)
	2	218.0(10)	214.5(10)	770.9(10)	1548.8(10)	5363.2(10)
	3	241.2(10)	390.3(10)	556.5(10)	1090.3(10)	5659.6(10)
GES-EM	1	6108.4(10)	10986.4(10)	27000.0(9)	41716.7(3)	-
	2	5239.4(10)	10623.3(9)	27361.7(9)	41551.8(3)	-
	3	5759.5(10)	9929.5(10)	16974.4(9)	43077.7(1)	-
GES-EM+T	1	2742.5(10)	4529.7(10)	16303.5(10)	26213.7(9)	-
	2	2446.3(10)	4856.0(10)	11321.9(10)	25160.3(10)	-
	3	2319.6(10)	4496.3(10)	12568.3(10)	20154.0(9)	-
SEM+RT	1	5.3(10)†	11.1(10)†	25.9(10)†	60.3(10)†	400.9(10)†
	2	6.6(10)†	12.3(9)†	39.8(10)†	94.5(10)†	650.9(10)†
	3	7.4(10)†	14.9(10)†	72.9(10)†	112.0(10)†	779.6(10)†

Tempo Alarm1 mar Tempo in secondi

Tempo HailFinder mar						
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	10.4(10)	5.8(10)	6.9(10)	11.5(10)	92.2(10)
MWST-EM	1	400.6(10)	620.4(10)	1368.2(10)	2671.1(10)	10359.3(10)
	2	365.6(10)	724.2(10)	1541.6(10)	2798.7(10)	11254.4(10)
	3	350.7(10)	634.4(10)	1503.4(10)	2727.3(10)	13071.8(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	<b>8.5(10)*</b>	<b>5.3(10)*</b>	<b>5.2(10)*</b>	<b>7.8(10)*</b>	<b>38.5(10)*</b>
	2	<b>8.8(10)*</b>	<b>5.4(10)*</b>	<b>6.2(10)*</b>	<b>9.2(10)*</b>	<b>44.2(10)*</b>
	3	<b>8.5(10)*</b>	<b>5.6(10)*</b>	<b>6.9(10)*</b>	<b>12.5(10)*</b>	<b>60.4(10)*</b>
k10+T	1	402.1(10)	621.9(10)	1370.1(10)	2674.1(10)	10381.5(10)
	2	367.2(10)	725.8(10)	1543.8(10)	2802.5(10)	11279.1(10)
	3	352.2(10)	636.1(10)	1506.1(10)	2731.6(10)	13098.3(10)
SEM	1	17.8(10)	45.8(10)	102.2(10)	243.1(10)	1302.7(10)
	2	22.8(10)	59.6(10)	144.0(10)	464.7(10)	1854.4(5)
	3	28.3(10)	55.3(10)	198.9(10)	594.1(10)	2291.7(8)
SEM+T	1	421.1(10)	664.7(10)	1550.6(10)	2933.7(10)	12238.0(10)
	2	412.6(10)	779.6(10)	1689.2(10)	3401.3(9)	13269.7(9)
	3	425.0(10)	740.2(10)	1866.5(10)	3341.2(10)	14858.6(7)
SEM.bic	1	21.0(10)	45.1(10)	109.8(10)	241.9(10)	1086.1(8)
	2	28.5(10)	50.7(10)	181.1(10)	406.3(10)	1543.5(6)
	3	38.2(10)	82.1(10)	201.5(10)	486.5(10)	2502.8(8)
SEM.bic+T	1	423.2(10)	666.9(10)	1523.4(10)	2930.8(10)	9950.1(9)
	2	413.2(10)	797.9(10)	1699.6(10)	3325.4(10)	12292.8(8)
	3	418.2(10)	724.8(10)	1907.0(10)	3318.1(10)	15335.4(6)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	17.3(10)†	54.9(10)†	101.9(10)†	271.4(10)†	1386.7(10)†
	2	31.2(10)†	77.6(10)†	161.8(10)†	487.3(10)†	2701.0(7)†
	3	36.4(10)†	72.4(10)†	222.7(10)†	608.0(10)†	2821.6(8)†

Tempo HailFinder mar Tempo in secondi

Tempo Hepar2 mar						
Metodo	p	100	200	500	1000	5000
MMHC	-	22.4(10)	38.9(10)	73.6(10)	382.9(10)	531.6(10)
MWST-EM	1	561.4(10)	839.9(10)	1693.7(10)	2261.4(10)	14550.9(10)
	2	441.3(10)	859.2(10)	2001.8(10)	2425.4(10)	13139.4(10)
	3	538.4(10)	874.9(10)	1488.0(10)	2322.0(10)	9652.4(6)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	22.0(10)*	<b>42.6(10)*</b>	<b>72.5(10)*</b>	352.0(10)*	<b>452.2(10)*</b>
	2	<b>25.3(10)*</b>	<b>43.0(10)*</b>	<b>119.8(10)*</b>	687.0(10)*	<b>2290.7(10)*</b>
	3	<b>26.2(10)*</b>	<b>57.0(10)*</b>	<b>139.6(10)*</b>	1002.8(10)*	2876.1(10)
k10+T	1	564.3(10)	842.8(10)	1697.1(10)	2265.9(10)	14577.9(10)
	2	444.2(10)	862.3(10)	2005.5(10)	2430.8(10)	13168.9(10)
	3	541.4(10)	878.3(10)	1492.1(10)	2328.1(10)	9686.1(6)
SEM	1	20.8(10)	43.9(10)	121.1(10)	<b>244.9(10)</b>	1730.8(10)
	2	25.3(10)	55.9(10)	199.0(10)	<b>398.1(10)</b>	3033.4(10)
	3	33.5(10)	79.4(10)	243.8(10)	<b>549.7(10)</b>	3772.3(10)
SEM+T	1	584.8(10)	902.5(10)	1857.8(10)	2460.5(10)	16087.3(10)
	2	473.1(10)	917.7(10)	2211.6(10)	2992.9(10)	16512.1(10)
	3	585.2(10)	965.8(10)	1718.3(10)	2946.2(10)	12611.0(6)
SEM.bic	1	<b>19.3(10)</b>	53.7(10)	137.5(10)	294.1(10)	1606.1(10)
	2	25.7(10)	72.8(10)	190.6(10)	438.8(10)	2911.6(10)
	3	37.4(10)	95.4(10)	236.6(10)	575.1(10)	<b>2771.8(10)</b>
SEM.bic+T	1	583.9(10)	944.6(10)	1878.2(10)	2553.1(10)	15950.5(10)
	2	467.6(10)	966.3(10)	2238.1(10)	2837.1(10)	15048.4(10)
	3	576.3(10)	993.5(10)	1820.1(10)	2772.8(10)	12545.6(6)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	27.2(10)†	49.8(10)†	119.9(10)†	267.7(10)†	1606.4(10)†
	2	30.5(10)†	65.4(10)†	228.5(10)†	473.4(10)†	3222.0(10)†
	3	35.7(10)†	71.6(10)†	289.4(10)†	523.9(10)†	3573.4(10)†

Tempo Hepar2 mar Tempo in secondi

		Tempo Asia nmar				
Metodo	p	100	200	500	1000	5000
MMHC	-	0.2(10)	0.2(10)	0.4(10)	0.9(10)	8.0(10)
MWST-EM	1	25.3(10)	38.5(10)	116.6(10)	287.0(10)	1423.1(10)
	2	31.3(10)	62.1(10)	118.2(10)	285.7(10)	1461.2(10)
	3	47.2(10)	68.1(10)	140.8(10)	317.7(10)	1775.0(10)
CC	1	0.1(10)	0.2(10)	0.3(10)	0.6(10)	4.9(10)
	2	0.1(10)	0.2(10)	0.4(10)	0.7(10)	4.4(10)
	3	0.2(10)	0.2(10)	0.4(10)	0.5(10)	3.2(10)
CC+T	1	25.4(10)	38.8(10)	117.0(10)	287.5(10)	1427.3(10)
	2	31.5(10)	62.4(10)	118.6(10)	286.7(10)	1464.6(10)
	3	47.4(10)	68.3(10)	141.1(10)	318.3(10)	1778.6(10)
AC	1	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
	2	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
	3	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>	<b>0.1(10)</b>
AC+T	1	25.3(10)	38.6(10)	116.7(10)	287.0(10)	1423.2(10)
	2	31.4(10)	62.2(10)	118.3(10)	285.8(10)	1461.3(10)
	3	47.3(10)	68.2(10)	140.8(10)	317.8(10)	1775.1(10)
k10	1	0.1(10)*	0.1(10)*	0.2(10)*	0.5(10)*	3.4(10)*
	2	0.1(10)*	0.2(10)*	0.2(10)*	0.6(10)*	4.4(10)*
	3	0.2(10)*	0.2(10)*	0.3(10)*	0.6(10)*	5.1(10)*
k10+T	1	25.4(10)	38.6(10)	116.7(10)	287.3(10)	1426.6(10)
	2	31.4(10)	62.3(10)	118.4(10)	286.1(10)	1466.3(10)
	3	47.3(10)	68.3(10)	141.0(10)	318.2(10)	1781.2(10)
SEM	1	0.1(10)	0.2(10)	0.2(10)	0.3(10)	0.4(10)
	2	0.2(10)	0.2(10)	0.2(10)	0.3(10)	0.4(10)
	3	0.2(10)	0.2(10)	0.3(10)	0.3(10)	0.4(10)
SEM+T	1	25.4(10)	38.7(10)	116.8(10)	287.3(10)	1423.8(10)
	2	31.5(10)	62.4(10)	118.4(10)	286.0(10)	1461.6(10)
	3	47.4(10)	68.4(10)	141.1(10)	317.9(10)	1775.5(10)
SEM.bic	1	0.1(10)	0.2(10)	0.2(10)	0.3(10)	0.4(10)
	2	0.2(10)	0.2(10)	0.2(10)	0.4(10)	0.5(10)
	3	0.2(10)	0.2(10)	0.2(10)	0.3(10)	0.5(10)
SEM.bic+T	1	25.4(10)	38.7(10)	116.8(10)	287.4(10)	1423.6(10)
	2	31.5(10)	62.4(10)	118.5(10)	286.1(10)	1461.6(10)
	3	47.4(10)	68.3(10)	141.0(10)	318.0(10)	1775.5(10)
GES-EM	1	61.6(10)	141.3(9)	434.0(9)	607.5(10)	3045.4(10)
	2	57.6(10)	126.4(10)	318.2(10)	687.5(10)	3214.7(10)
	3	51.7(10)	126.4(10)	322.2(10)	685.1(10)	3403.0(10)
GES-EM+T	1	83.0(10)	212.1(10)	605.6(10)	1106.0(10)	5057.8(10)
	2	102.9(7)	226.4(10)	599.6(10)	1424.6(10)	5071.4(10)
	3	131.0(10)	229.5(10)	620.7(10)	1307.3(10)	5871.5(10)
SEM+RT	1	0.1(10)†	0.2(10)†	0.2(10)†	0.2(10)†	0.3(10)†
	2	0.2(10)†	0.2(10)†	0.2(10)†	0.3(10)†	0.5(10)†
	3	0.2(10)†	0.2(10)†	0.2(10)†	0.3(10)†	0.5(10)†

Tempo Asia nmar Tempo in secondi

Tempo Child nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	0.7(10)	0.6(10)	1.3(10)	2.7(10)	38.0(10)
MWST-EM	1	147.2(10)	263.5(10)	488.0(10)	1313.3(10)	4691.8(10)
	2	114.3(10)	213.0(10)	622.5(10)	1004.1(10)	3872.0(10)
	3	108.3(10)	257.2(10)	501.7(10)	970.1(10)	3998.7(10)
CC	1	1.4(10)	1.5(10)	0.7(10)	1.2(10)	4.9(10)
	2	3.2(10)	1.9(10)	1.4(10)	1.1(10)	4.9(10)
	3	15.6(9)	3.5(10)	2.7(10)	1.6(10)	4.0(10)
CC+T	1	147.4(10)	263.8(10)	488.4(10)	1314.1(10)	4701.3(10)
	2	114.6(10)	213.2(10)	622.9(10)	1004.8(10)	3876.0(10)
	3	109.0(9)	257.5(10)	502.1(10)	970.8(10)	4002.4(10)
AC	1	0.6(10)	0.6(10)	<b>0.5(10)</b>	<b>0.8(10)</b>	<b>4.0(10)</b>
	2	1.4(10)	0.7(10)	<b>0.8(10)</b>	<b>0.8(10)</b>	<b>3.6(10)</b>
	3	1.0(10)	0.7(10)	<b>0.6(10)</b>	<b>0.8(10)</b>	<b>3.6(10)</b>
AC+T	1	147.3(10)	263.6(10)	488.1(10)	1313.4(10)	4692.0(10)
	2	114.5(10)	213.1(10)	622.7(10)	1004.2(10)	3872.2(10)
	3	108.4(10)	257.4(10)	501.8(10)	970.2(10)	3998.9(10)
k10	1	<b>0.4(10)*</b>	<b>0.5(10)*</b>	0.8(10)*	1.4(10)*	19.0(10)*
	2	<b>0.5(10)*</b>	<b>0.5(10)*</b>	1.0(10)*	2.1(10)*	20.4(10)*
	3	<b>0.5(10)*</b>	<b>0.6(10)*</b>	1.2(10)*	2.5(10)*	26.6(10)*
k10+T	1	147.4(10)	263.8(10)	488.4(10)	1314.1(10)	4700.8(10)
	2	114.5(10)	213.2(10)	623.0(10)	1005.1(10)	3883.6(10)
	3	108.5(10)	257.5(10)	502.3(10)	971.3(10)	4010.9(10)
SEM	1	1.0(10)	2.7(10)	5.0(10)	12.3(10)	78.6(10)
	2	0.9(10)	3.1(10)	6.7(10)	17.0(10)	90.1(10)
	3	1.3(10)	3.9(10)	8.1(10)	17.7(10)	93.2(10)
SEM+T	1	149.9(10)	266.9(10)	492.9(10)	1325.5(10)	4784.5(10)
	2	115.7(10)	217.0(10)	633.5(10)	1026.6(10)	3942.8(10)
	3	109.5(10)	260.0(10)	509.5(10)	999.1(10)	4124.2(10)
SEM.bic	1	1.2(10)	3.0(10)	5.9(10)	9.5(10)	76.6(10)
	2	1.4(10)	3.0(10)	7.1(10)	16.0(10)	99.1(10)
	3	0.9(10)	2.6(10)	8.4(10)	26.3(10)	95.9(10)
SEM.bic+T	1	148.2(10)	266.2(10)	502.2(10)	1332.9(10)	4804.3(10)
	2	116.4(10)	216.1(10)	629.8(10)	1020.0(10)	3966.7(10)
	3	110.4(10)	259.7(10)	513.5(10)	991.5(10)	4152.4(10)
GES-EM	1	516.1(10)	1127.0(10)	3564.1(10)	8805.4(10)	37846.1(9)
	2	686.3(10)	1156.7(10)	3272.6(10)	6761.4(10)	37064.0(8)
	3	359.0(10)	1010.4(10)	2834.0(10)	6038.0(10)	35816.0(8)
GES-EM+T	1	631.5(10)	1429.4(10)	3244.6(9)	7443.2(10)	41148.4(9)
	2	609.8(10)	1094.6(9)	3222.1(10)	6648.8(4)	38128.8(4)
	3	-	-	-	-	37114.9(7)
SEM+RT	1	1.6(10)†	3.3(10)†	4.0(10)†	12.2(10)†	75.0(10)†
	2	1.3(10)†	3.3(10)†	6.2(10)†	18.6(10)†	100.4(10)†
	3	1.5(10)†	3.5(10)†	8.2(10)†	18.0(10)†	120.1(10)†

Tempo Child nmar Tempo in secondi

Tempo Insurance nmar						
Metodo	p	100	200	500	1000	5000
MMHC	-	1.6(10)	1.4(10)	2.2(10)	4.1(10)	53.7(10)
MWST-EM	1	148.6(10)	267.4(10)	695.2(10)	1320.2(10)	7070.5(10)
	2	184.6(10)	347.3(10)	800.9(10)	1617.3(10)	6598.8(10)
	3	192.1(10)	285.4(10)	823.8(10)	1654.6(10)	6670.4(10)
CC	1	26.8(10)	1.9(10)	1.9(10)	2.1(10)	11.7(10)
	2	813.7(7)	7.0(9)	5.8(10)	2.6(10)	6.4(10)
	3	6253.8(4)	1327.1(7)	221.2(8)	4.8(8)	5.8(10)
CC+T	1	149.2(10)	268.1(10)	696.4(10)	1321.4(10)	7079.2(10)
	2	184.9(10)	347.9(10)	801.6(10)	1619.3(10)	6604.2(10)
	3	191.5(9)	286.1(10)	824.7(10)	1654.9(9)	6674.4(10)
AC	1	3.0(10)	<b>1.0(10)</b>	<b>0.9(10)</b>	<b>1.3(10)</b>	<b>4.6(10)</b>
	2	33.6(10)	1.5(10)	<b>1.4(10)</b>	<b>1.3(10)</b>	<b>4.0(10)</b>
	3	53.0(10)	2.5(10)	<b>1.1(10)</b>	<b>1.3(10)</b>	<b>4.2(10)</b>
AC+T	1	148.9(10)	267.7(10)	695.5(10)	1320.6(10)	7070.9(10)
	2	184.9(10)	347.6(10)	801.3(10)	1617.7(10)	6599.2(10)
	3	192.5(10)	285.8(10)	824.1(10)	1655.1(10)	6670.8(10)
k10	1	<b>1.3(10)*</b>	1.2(10)*	1.6(10)*	3.5(10)*	23.4(10)*
	2	<b>1.3(10)*</b>	<b>1.3(10)*</b>	1.7(10)*	3.7(10)*	32.3(10)*
	3	<b>1.3(10)*</b>	<b>1.3(10)*</b>	1.9(10)*	4.3(10)*	51.7(10)*
k10+T	1	149.0(10)	267.8(10)	695.8(10)	1321.4(10)	7081.9(10)
	2	185.0(10)	347.9(10)	801.7(10)	1618.9(10)	6612.4(10)
	3	192.6(10)	285.9(10)	824.9(10)	1656.3(10)	6685.3(10)
SEM	1	2.3(10)	4.9(10)	13.3(10)	32.1(10)	245.7(10)
	2	3.6(10)	5.8(10)	17.6(10)	55.8(10)	324.4(10)
	3	3.5(10)	9.1(10)	18.8(10)	56.4(10)	448.2(10)
SEM+T	1	152.7(10)	271.4(10)	707.7(10)	1358.1(10)	7274.0(10)
	2	190.8(10)	351.3(10)	818.3(10)	1686.0(10)	7123.2(8)
	3	197.0(10)	295.8(10)	840.8(10)	1710.0(10)	7165.3(10)
SEM.bic	1	2.1(10)	4.0(10)	11.2(10)	23.9(10)	267.5(10)
	2	4.4(10)	6.4(10)	17.9(10)	37.9(10)	257.5(10)
	3	2.9(10)	6.5(10)	21.8(10)	32.7(10)	329.4(10)
SEM.bic+T	1	151.9(10)	272.6(10)	709.0(10)	1349.3(10)	7253.0(10)
	2	190.8(10)	351.8(10)	815.7(10)	1655.1(10)	6881.9(10)
	3	195.1(10)	291.5(10)	843.6(10)	1681.2(10)	7042.7(10)
GES-EM	1	1451.5(10)	3360.9(10)	9405.9(10)	25618.3(7)	-
	2	1492.0(10)	3413.7(9)	11907.7(9)	21316.8(8)	-
	3	2808.4(10)	5866.7(10)	11483.3(10)	20362.5(10)	-
GES-EM+T	1	1103.3(10)	2173.8(9)	6439.6(10)	13607.1(10)	-
	2	1077.3(10)	2333.3(9)	6468.8(10)	13642.7(10)	-
	3	1027.0(9)	1943.8(10)	5965.3(9)	13929.0(10)	-
SEM+RT	1	3.5(10)†	5.5(10)†	14.8(10)†	29.1(10)†	275.9(10)†
	2	3.9(10)†	7.1(10)†	18.3(10)†	42.3(10)†	410.9(10)†
	3	4.1(10)†	8.9(10)†	25.9(10)†	55.1(10)†	435.7(10)†

Tempo Insurance nmar Tempo in secondi

Metodo	p	Tempo Alarm1 nmar				
		100	200	500	1000	5000
MMHC	-	4.4(10)	4.1(10)	6.0(10)	8.0(10)	60.0(10)
MWST-EM	1	220.5(10)	379.0(10)	952.0(10)	928.4(10)	4814.9(10)
	2	209.7(10)	387.6(10)	480.2(10)	1010.9(10)	8989.7(10)
	3	251.7(10)	315.5(10)	789.8(10)	1019.3(10)	7199.2(10)
CC	1	6927.9(10)	356.9(9)	61.8(9)	6.9(10)	9.8(10)
	2	398.8(5)	6197.2(4)	5366.9(5)	12233.5(6)	8.7(10)
	3	16.1(4)	11.4(2)	3947.7(2)	20.2(5)	714.0(7)
CC+T	1	221.2(10)	379.7(10)	952.9(10)	929.8(10)	4819.8(10)
	2	210.1(7)	388.2(8)	481.0(10)	1012.1(10)	8994.4(10)
	3	261.7(4)	193.7(2)	471.2(2)	1008.1(5)	7487.9(9)
AC	1	15.2(10)	<b>3.2(10)</b>	<b>3.4(10)</b>	<b>3.4(10)</b>	<b>8.2(10)</b>
	2	887.7(10)	5.3(10)	<b>4.3(10)</b>	<b>3.9(10)</b>	<b>7.8(10)</b>
	3	433.3(10)	5.3(10)	<b>4.4(10)</b>	<b>4.2(10)</b>	<b>7.7(10)</b>
AC+T	1	221.6(10)	379.5(10)	952.3(10)	928.9(10)	4815.5(10)
	2	210.2(10)	388.0(10)	480.6(10)	1011.3(10)	8990.3(10)
	3	252.1(10)	315.9(10)	790.2(10)	1019.8(10)	7199.8(10)
k10	1	5.9(10)*	5.5(10)*	5.5(10)*	6.7(10)*	31.6(10)*
	2	<b>4.2(10)*</b>	<b>4.5(10)*</b>	6.5(10)*	9.1(10)*	36.8(10)*
	3	<b>4.4(10)*</b>	<b>4.8(10)*</b>	7.5(10)*	12.0(10)*	49.5(10)*
k10+T	1	221.1(10)	379.6(10)	952.9(10)	931.6(10)	4831.9(10)
	2	210.4(10)	388.2(10)	481.3(10)	1012.9(10)	9007.3(10)
	3	252.3(10)	316.3(10)	791.2(10)	1022.0(10)	7218.5(10)
SEM	1	<b>5.3(10)</b>	10.4(10)	36.3(10)	79.6(10)	376.7(10)
	2	7.9(10)	13.2(10)	46.2(10)	92.2(10)	603.6(10)
	3	8.1(10)	13.7(10)	56.4(10)	103.3(10)	498.7(10)
SEM+T	1	227.4(10)	390.9(10)	996.3(10)	1012.4(10)	5239.6(10)
	2	220.8(10)	404.7(10)	526.9(10)	1216.3(10)	9680.3(10)
	3	258.5(10)	338.2(10)	855.3(10)	1182.2(10)	7831.7(10)
SEM.bic	1	5.9(10)	10.9(10)	34.7(10)	51.3(10)	377.1(10)
	2	7.0(10)	12.4(10)	46.9(10)	87.9(10)	518.1(10)
	3	9.3(10)	12.2(10)	55.2(10)	94.8(10)	513.1(10)
SEM.bic+T	1	225.3(10)	386.9(10)	992.8(10)	997.4(10)	5189.2(10)
	2	217.3(10)	400.5(10)	557.3(10)	1111.1(10)	9529.1(10)
	3	257.6(10)	335.0(10)	859.5(10)	1199.1(10)	7912.6(10)
GES-EM	1	7304.6(8)	16057.3(9)	28447.9(10)	35886.7(3)	-
	2	5901.6(9)	11325.4(10)	21425.9(10)	-	-
	3	8586.0(9)	10046.9(9)	20390.3(10)	35636.2(10)	-
GES-EM+T	1	2541.2(10)	5019.7(10)	11475.3(10)	27460.2(9)	-
	2	3020.8(10)	4393.8(10)	10375.6(10)	23680.4(10)	-
	3	2304.7(10)	4630.1(10)	10814.8(10)	27481.7(9)	-
SEM+RT	1	5.6(10)†	10.9(10)†	33.5(10)†	70.6(10)†	429.9(10)†
	2	8.9(10)†	15.3(10)†	44.8(10)†	101.2(10)†	764.6(10)†
	3	10.7(10)†	17.6(10)†	61.2(10)†	137.5(10)†	748.8(10)†

Tempo Alarm1 nmar Tempo in secondi

		Tempo HailFinder nmar				
Metodo	p	100	200	500	1000	5000
<i>MMHC</i>	-	10.4(10)	5.8(10)	6.9(10)	11.5(10)	92.2(10)
MWST-EM	1	410.9(10)	752.3(10)	1437.7(10)	2496.0(10)	10047.9(10)
	2	469.6(10)	661.1(10)	1435.4(10)	2716.9(10)	13087.0(10)
	3	393.2(10)	621.1(10)	1435.9(10)	2748.9(10)	13326.2(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	<b>7.7(10)*</b>	<b>5.1(10)*</b>	<b>5.5(10)*</b>	<b>7.1(10)*</b>	<b>39.5(10)*</b>
	2	<b>7.5(10)*</b>	<b>5.5(10)*</b>	<b>6.2(10)*</b>	<b>8.5(10)*</b>	<b>47.3(10)</b>
	3	<b>8.6(10)*</b>	<b>5.8(10)*</b>	<b>7.0(10)*</b>	<b>10.8(10)*</b>	<b>61.3(10)*</b>
k10+T	1	412.3(10)	753.8(10)	1439.6(10)	2499.2(10)	10070.6(10)
	2	471.1(10)	663.0(10)	1437.5(10)	2720.3(10)	13112.1(10)
	3	394.9(10)	623.0(10)	1438.3(10)	2752.8(10)	13353.1(10)
SEM	1	17.5(10)	57.1(10)	119.1(10)	229.3(10)	1400.4(7)
	2	21.5(10)	87.4(10)	184.5(10)	391.7(10)	2160.9(5)
	3	26.5(10)	73.3(10)	315.3(10)	413.8(10)	2244.3(7)
SEM+T	1	441.6(10)	802.8(10)	1593.7(10)	2832.5(10)	12764.4(7)
	2	515.3(10)	749.3(10)	1792.4(10)	3254.2(10)	14973.7(5)
	3	448.7(10)	798.6(10)	1832.8(10)	3425.7(10)	16047.8(6)
SEM.bic	1	19.7(10)	48.0(10)	118.2(10)	222.5(10)	1315.0(8)
	2	42.9(10)	94.0(10)	213.6(10)	393.9(10)	1513.4(5)
	3	28.5(10)	77.0(10)	347.7(10)	480.4(10)	1846.7(8)
SEM.bic+T	1	432.2(10)	796.6(10)	1567.5(10)	2850.1(10)	13121.7(8)
	2	522.2(10)	755.0(10)	1704.0(10)	3143.7(10)	15137.0(5)
	3	463.8(10)	782.0(10)	1870.4(10)	3258.6(10)	14733.3(7)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	20.4(10)†	54.6(10)†	120.8(10)†	326.7(10)†	2641.2(9)†
	2	30.2(10)†	103.1(10)†	204.6(10)†	399.6(10)†	2141.6(8)
	3	29.0(10)†	96.3(10)†	260.2(10)†	472.7(10)†	3329.8(7)†

Tempo HailFinder nmar Tempo in secondi

		Tempo Hepar2 nmar				
Metodo	p	100	200	500	1000	5000
MMHC	-	22.4(10)	38.9(10)	73.6(10)	382.9(10)	531.6(10)
MWST-EM	1	435.4(10)	780.1(10)	1901.9(10)	3300.1(10)	13024.5(10)
	2	528.3(10)	855.7(10)	1907.8(10)	3595.6(10)	13167.1(10)
	3	424.2(10)	776.9(10)	1601.4(10)	2329.7(10)	8964.7(10)
CC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
CC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
AC+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
k10	1	<b>22.2(10)*</b>	49.4(10)*	<b>125.7(10)*</b>	683.5(10)*	2181.9(10)
	2	<b>25.0(10)*</b>	<b>48.4(10)*</b>	<b>116.3(10)*</b>	930.8(10)*	3071.7(9)*
	3	26.5(10)*	<b>57.6(10)*</b>	<b>205.7(10)*</b>	928.4(10)*	2952.0(9)*
k10+T	1	438.4(10)	783.0(10)	1905.5(10)	3304.5(10)	13052.8(10)
	2	531.3(10)	858.9(10)	1911.6(10)	3601.0(10)	13196.3(10)
	3	427.1(10)	780.2(10)	1605.7(10)	2336.3(10)	8996.6(10)
SEM	1	23.3(10)	51.2(10)	172.7(10)	325.7(10)	2057.3(10)
	2	34.0(10)	73.6(10)	242.9(10)	519.4(10)	2469.1(10)
	3	36.4(10)	91.1(10)	297.0(10)	<b>534.9(10)</b>	<b>2856.4(10)</b>
SEM+T	1	467.8(10)	839.1(10)	2070.3(10)	3523.2(10)	14868.5(10)
	2	556.5(10)	929.0(10)	2140.2(10)	4022.0(10)	15453.8(10)
	3	486.7(10)	871.6(10)	1954.7(10)	3967.8(10)	14642.3(10)
SEM.bic	1	26.8(10)	<b>49.3(10)</b>	152.3(10)	<b>310.4(10)</b>	<b>1722.0(10)</b>
	2	32.1(10)	69.8(10)	219.1(10)	<b>475.2(10)</b>	<b>2351.5(10)</b>
	3	<b>23.7(10)</b>	87.7(10)	247.0(10)	578.4(10)	3050.0(10)
SEM.bic+T	1	477.2(10)	828.8(10)	2064.5(10)	3575.9(10)	14752.0(10)
	2	563.8(10)	951.6(10)	2130.8(10)	3986.3(10)	15404.4(10)
	3	472.7(10)	882.9(10)	2137.2(10)	3013.2(10)	13469.8(10)
GES-EM	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
GES-EM+T	1	-	-	-	-	-
	2	-	-	-	-	-
	3	-	-	-	-	-
SEM+RT	1	27.7(10)†	65.8(10)†	163.7(10)†	274.7(10)†	1713.8(10)†
	2	32.9(10)†	80.9(10)†	218.8(10)†	551.2(10)†	2211.2(10)†
	3	44.1(10)†	73.2(10)†	330.2(10)†	829.9(10)†	3234.3(10)†

Tempo Hepar2 nmar Tempo in secondi

## AUC k-nn e SEM+T

AUC ensemble learning Asia mcar						
Metodo	p	100	200	500	1000	5000
k10	1	<b>0.4677(10)</b>	0.5866(10)	0.5866(10)	0.6244(10)	0.7689(10)
	2	0.4365(10)	0.4765(10)	0.6431(10)	0.6154(10)	0.7914(10)
	3	0.3864(10)	0.4023(10)	0.5868(10)	0.5430(10)	0.7729(10)
SEM+T	1	<b>0.4677(10)</b>	0.6408(10)	<b>0.8871(10)*</b>	<b>0.9247(10)*</b>	<b>0.9247(10)</b>
	2	<b>0.4677(10)</b>	<b>0.6884(10)*</b>	<b>0.8315(10)</b>	<b>0.8781(10)*</b>	<b>0.9247(10)</b>
	3	<b>0.5698(10)</b>	<b>0.5866(10)</b>	0.7208(10)	0.7294(10)	<b>0.9247(10)</b>
k10 + SEM+T	1	0.4553(10)	<b>0.6619(10)</b>	0.8008(10)	0.8637(10)	0.9123(10)
	2	0.4373(10)	0.5742(10)	0.8180(10)	0.8283(10)	0.9032(10)
	3	0.4183(10)∇	0.4693(10)	<b>0.7276(10)</b>	<b>0.7608(10)</b>	0.9099(10)
SEM+RT	1	<i>0.5282(10)</i>	<i>0.6660(10)</i>	<i>0.8327(10)</i>	<i>0.8315(10)</i>	<i>0.9019(10)</i>
	2	<i>0.4300(10)</i>	<i>0.6858(10)</i>	<i>0.8781(10)</i>	<i>0.8315(10)</i>	<i>0.9247(10)</i>
	3	<i>0.3279(10)</i>	<i>0.5554(10)</i>	<i>0.7404(10)</i>	<i>0.6660(10)</i>	<i>0.9247(10)</i>
k10 + SEM+RT	1	<i>0.5127(10)</i>	<i>0.6327(10)</i>	<i>0.7778(10)</i>	<i>0.8271(10)</i>	<i>0.8937(10)</i>
	2	<i>0.4115(10)</i>	<i>0.5981(10)</i>	<i>0.8679(10)</i>	<i>0.8145(10)</i>	<i>0.9111(10)</i>
	3	<i>0.3404(10)</i>	<i>0.4483(10)</i>	<i>0.7317(10)</i>	<i>0.7068(10)</i>	<i>0.8994(10)</i>

**AUC Asia mcar** In grassetto i risultati migliori.

AUC ensemble learning Asia mar						
Metodo	p	100	200	500	1000	5000
k10	1	0.4865(10)	0.5803(10)	0.5930(10)	0.6908(10)	0.7464(10)
	2	0.5282(10)	0.5866(10)	0.5241(10)	0.6772(10)	0.8168(10)
	3	0.5470(10)	0.5241(10)	0.6090(10)	0.5380(10)	0.6152(10)
SEM+T	1	<b>0.5782(10)</b>	<b>0.7156(10)</b>	<b>0.8781(10)</b>	<b>0.8781(10)</b>	<b>0.9247(10)</b>
	2	<b>0.6090(10)</b>	<b>0.6408(10)</b>	<b>0.8781(10)</b>	<b>0.8315(10)</b>	<b>0.9247(10)</b>
	3	0.5004(10)	<b>0.5866(10)</b>	<b>0.7859(10)</b>	<b>0.8315(10)</b>	0.8377(10)
k10 + SEM+T	1	0.5354(10)	0.6474(10)	0.8107(10)	0.8731(10)	0.8813(10)
	2	0.6025(10)	0.6259(10)	0.8064(10)	0.8123(10)	0.8913(10)
	3	<b>0.5545(10)</b>	0.5416(10)	0.7723(10)	0.6785(10)	<b>0.8488(10)</b>
SEM+RT	1	0.6006(10)	0.5965(10)	0.9247(10)	0.8315(10)	0.9247(10)
	2	0.6006(10)	0.6951(10)	0.6858(10)	0.8315(10)	0.9247(10)
	3	0.4489(10)	0.4865(10)	0.6927(10)	0.6451(10)	0.6979(10)
k10 + SEM+RT	1	0.5235(10)	0.6079(10)	0.8409(10)	0.8731(10)	0.8915(10)
	2	0.5433(10)	0.6779(10)	0.6733(10)	0.8538(10)	0.8924(10)
	3	0.5234(10)	0.4828(10)	0.7723(10)	0.6596(10)	0.8341(10)

**AUC Asia mar** In grassetto i risultati migliori.

AUC ensemble learning Asia nmar						
Metodo	P	100	200	500	1000	5000
k10	1	<b>0.5698(10)</b>	0.4845(10)	0.6056(10)	0.5866(10)	0.6994(10)
	2	0.5803(10)	0.4927(10)	0.6387(10)	0.6027(10)	0.6564(10)
	3	0.4905(10)	0.4365(10)	0.5519(10)	0.6660(10)	0.7165(10)
SEM+T	1	0.5248(10)	<b>0.6382(10)</b>	<b>0.8092(10)</b>	<b>0.8781(10)*</b>	<b>0.8330(10)</b>
	2	<b>0.6365(10)</b>	<b>0.6006(10)</b>	<b>0.8315(10)*</b>	<b>0.8315(10)*</b>	<b>0.9247(10)</b>
	3	0.4865(10)†	<b>0.5782(10)</b>	0.6590(10)	0.7136(10)	<b>0.8781(10)</b>
k10 + SEM+T	1	0.5537(10)	0.5909(10)	0.7961(10)	0.8390(10)	0.8220(10)
	2	0.6196(10)	0.5464(10)	0.8136(10)	0.8168(10)	0.8740(10)
	3	<b>0.5441(10)∇</b>	0.5222(10)	<b>0.7610(10)</b>	<b>0.7325(10)</b>	0.8652(10)
SEM+RT	1	<i>0.6313(10)</i>	<i>0.6660(10)</i>	<i>0.6927(10)</i>	<i>0.8315(10)</i>	<i>0.7871(10)</i>
	2	<i>0.4122(10)</i>	<i>0.6006(10)</i>	<i>0.8093(10)</i>	<i>0.8315(10)</i>	<i>0.9247(10)</i>
	3	<i>0.4112(10)</i>	<i>0.4914(10)</i>	<i>0.6215(10)</i>	<i>0.6858(10)</i>	<i>0.8093(10)</i>
k10 + SEM+RT	1	<i>0.6014(10)</i>	<i>0.5891(10)</i>	<i>0.7280(10)</i>	<i>0.7495(10)</i>	<i>0.8022(10)</i>
	2	<i>0.4922(10)</i>	<i>0.5464(10)</i>	<i>0.7918(10)</i>	<i>0.7968(10)</i>	<i>0.8736(10)</i>
	3	<i>0.4092(10)</i>	<i>0.5004(10)</i>	<i>0.6636(10)</i>	<i>0.6956(10)</i>	<i>0.8234(10)</i>

**AUC Asia nmar** In grassetto i risultati migliori.

AUC ensemble learning Child mcar						
Metodo	p	100	200	500	1000	5000
k10	1	0.5626(10)	<b>0.6738(10)</b>	<b>0.7772(10)</b>	<b>0.8767(10)</b>	<b>1.0000(10)</b>
	2	<b>0.5094(10)</b>	<b>0.6444(10)*</b>	<b>0.7602(10)*</b>	0.7836(10)	<b>1.0000(10)*</b>
	3	<b>0.4634(10)</b>	<b>0.5991(10)*</b>	<b>0.6782(10)*</b>	<b>0.7393(10)</b>	<b>0.9377(10)</b>
SEM+T	1	0.6058(10)	0.6126(10)	0.6856(10)	0.8050(10)	0.9614(10)
	2	0.4797(10)	0.5187(10)	0.5722(10)	0.7425(10)	0.7988(10)
	3	0.4010(10)	0.4241(10)	0.4785(10)	0.5789(10)	0.8014(9)
k10 + SEM+T	1	<b>0.6073(10)</b>	0.6652(10)	0.7655(10)	0.8636(10)	0.9993(10)
	2	0.4946(10)	0.5940(10)	0.7051(10)	<b>0.7913(10)</b>	0.9746(10)
	3	0.4405(10)	0.5436(10)	0.6100(10)	0.7075(10)	0.9310(9)
SEM+RT	1	<i>0.5776(10)</i>	<i>0.6005(10)</i>	<i>0.7001(10)</i>	<i>0.8102(10)</i>	<i>0.8839(10)</i>
	2	<i>0.5249(10)</i>	<i>0.5542(10)</i>	<i>0.6126(10)</i>	<i>0.6608(10)</i>	<i>0.9099(10)†</i>
	3	<i>0.4779(10)†</i>	<i>0.4984(10)†</i>	<i>0.5117(10)</i>	<i>0.5357(10)</i>	<i>0.8167(10)</i>
k10 + SEM+RT	1	<i>0.5907(10)</i>	<i>0.6603(10)</i>	<i>0.7479(10)</i>	<i>0.8721(10)</i>	<i>0.9940(10)</i>
	2	<i>0.5228(10)</i>	<i>0.6134(10)</i>	<i>0.7105(10)</i>	<i>0.7920(10)</i>	<i>0.9960(10)∇</i>
	3	<i>0.4922(10)∇</i>	<i>0.5711(10)</i>	<i>0.6206(10)</i>	<i>0.6698(10)</i>	<i>0.9361(10)</i>

**AUC Child mcar** In grassetto i risultati migliori.

AUC ensemble learning Child mar						
Metodo	P	100	200	500	1000	5000
k10	1	0.5633(10)	<b>0.6750(10)</b>	<b>0.7602(10)</b>	<b>0.8453(10)</b>	<b>1.0000(10)</b>
	2	0.5076(10)	<b>0.6406(10)</b>	<b>0.6968(10)*</b>	<b>0.7810(10)</b>	<b>0.9868(10)</b>
	3	0.4520(10)	<b>0.6464(10)</b>	<b>0.6870(10)</b>	<b>0.7763(10)</b>	<b>0.9367(10)</b>
SEM+T	1	0.5363(10)	0.6360(10)	0.6789(10)	0.7448(10)	0.9549(10)
	2	<b>0.5573(10)</b>	0.5603(10)	0.6136(10)	0.7033(10)	0.8326(10)
	3	0.4629(10)	0.5585(10)	0.6156(10)†	0.5637(10)	0.7167(10)
k10 + SEM+T	1	<b>0.5711(10)</b>	0.6707(10)	0.7410(10)	0.8290(10)	0.9990(10)
	2	0.5570(10)	0.6378(10)	0.6676(10)	0.7763(10)	0.9848(10)
	3	<b>0.4822(10)</b>	0.6406(10)	0.6741(10)	0.7546(10)	0.9221(10)
SEM+RT	1	<i>0.5201(10)</i>	<i>0.6249(10)</i>	<i>0.6783(10)</i>	<i>0.6988(10)</i>	<i>0.9491(10)</i>
	2	<i>0.5027(10)</i>	<i>0.6052(10)†</i>	<i>0.5644(10)</i>	<i>0.6676(10)</i>	<i>0.9291(10)</i>
	3	<i>0.4725(10)</i>	<i>0.5838(10)</i>	<i>0.5523(10)</i>	<i>0.6439(10)</i>	<i>0.7862(10)</i>
k10 + SEM+RT	1	<i>0.5667(10)</i>	<i>0.6705(10)</i>	<i>0.7421(10)</i>	<i>0.8340(10)</i>	<i>0.9983(10)</i>
	2	<i>0.5528(10)</i>	<i>0.6453(10)∇</i>	<i>0.6555(10)</i>	<i>0.7937(10)</i>	<i>0.9938(10)</i>
	3	<i>0.4968(10)</i>	<i>0.6613(10)</i>	<i>0.6325(10)</i>	<i>0.7746(10)</i>	<i>0.9357(10)</i>

**AUC Child mar** In grassetto i risultati migliori.

AUC ensemble learning Child nmar						
Metodo	p	100	200	500	1000	5000
k10	1	0.5719(10)	<b>0.6702(10)</b>	<b>0.7602(10)</b>	0.7822(10)	0.9684(10)
	2	<b>0.5821(10)</b>	<b>0.6396(10)</b>	<b>0.7294(10)*</b>	0.7612(10)	<b>0.9744(10)</b>
	3	0.5246(10)	<b>0.6568(10)</b>	<b>0.6739(10)*</b>	0.7303(10)	<b>0.9449(10)</b>
SEM+T	1	0.5566(10)	0.5879(10)	0.5583(10)	0.7504(10)	0.9357(10)
	2	0.5130(10)	0.5353(10)	0.5720(10)	0.7229(10)	0.8482(10)
	3	0.5101(10)	0.5583(10)	0.5141(10)	0.6754(10)	0.7185(10)
k10 + SEM+T	1	<b>0.6192(10)</b>	0.6403(10)	0.7075(10)	<b>0.7945(10)</b>	<b>0.9792(10)</b>
	2	0.5718(10)	0.6051(10)	0.6635(10)	<b>0.7904(10)</b>	0.9706(10)
	3	<b>0.5440(10)</b>	0.6329(10)	0.6381(10)	<b>0.7626(10)</b>	0.9406(10)
SEM+RT	1	<i>0.5373(10)</i>	<i>0.5987(10)</i>	<i>0.6478(10)</i>	<i>0.7523(10)</i>	<i>0.9357(10)</i>
	2	<i>0.5515(10)</i>	<i>0.5798(10)</i>	<i>0.5733(10)</i>	<i>0.6505(10)</i>	<i>0.9487(10)†</i>
	3	<i>0.4786(10)</i>	<i>0.5464(10)</i>	<i>0.5415(10)</i>	<i>0.6179(10)</i>	<i>0.7883(10)</i>
k10 + SEM+RT	1	<i>0.5709(10)</i>	<i>0.6475(10)</i>	<i>0.7324(10)</i>	<i>0.8077(10)</i>	<i>0.9822(10)</i>
	2	<i>0.5720(10)</i>	<i>0.6185(10)</i>	<i>0.7012(10)</i>	<i>0.7670(10)</i>	<i>0.9962(10)</i>
	3	<i>0.5263(10)</i>	<i>0.6498(10)</i>	<i>0.6523(10)</i>	<i>0.7771(10)</i>	<i>0.9627(10)</i>

**AUC Child nmar** In grassetto i risultati migliori.

AUC ensemble learning Insurance mcar						
Metodo	p	100	200	500	1000	5000
k10	1	<b>0.4902(10)</b>	0.5321(10)	<b>0.6137(10)*</b>	<b>0.6669(10)</b>	<b>0.7684(10)</b>
	2	<b>0.4195(10)</b>	<b>0.5295(10)</b>	<b>0.5422(10)</b>	<b>0.6045(10)</b>	0.6847(10)
	3	<b>0.3615(10)</b>	<b>0.4402(10)</b>	<b>0.4947(10)</b>	<b>0.5685(10)</b>	0.6235(10)
SEM+T	1	0.4343(10)	0.5037(10)	0.4901(10)	0.5791(10)	0.6892(10)
	2	0.3723(10)	0.4283(10)	0.4661(10)	0.4743(10)	0.6386(7)
	3	0.3497(10)	0.4007(10)	0.4244(10)	0.4604(10)	0.5945(1)
k10 + SEM+T	1	0.4726(10)	<b>0.5415(10)</b>	0.5957(10)	0.6546(10)	0.7661(10)
	2	0.3903(10)	0.4837(10)	0.5225(10)	0.6044(10)	<b>0.7033(7)</b>
	3	0.3419(10)	0.4354(10)	0.4666(10)	0.5305(10)	<b>0.7141(1)</b>
SEM+RT	1	<i>0.4137(10)</i>	<i>0.4579(10)</i>	<i>0.5291(10)</i>	<i>0.5874(10)</i>	<i>0.6849(10)</i>
	2	<i>0.3292(10)</i>	<i>0.4272(10)</i>	<i>0.4991(10)</i>	<i>0.4704(10)</i>	<i>0.5926(7)</i>
	3	<i>0.3209(10)</i>	<i>0.4080(10)</i>	<i>0.4127(10)</i>	<i>0.3837(9)</i>	<i>0.5118(2)</i>
k10 + SEM+RT	1	<i>0.4709(10)</i>	<i>0.5164(10)</i>	<i>0.6291(10)</i>	<i>0.6552(10)</i>	<i>0.7974(10)</i>
	2	<i>0.3739(10)</i>	<i>0.4722(10)</i>	<i>0.5298(10)</i>	<i>0.5811(10)</i>	<i>0.7050(7)</i>
	3	<i>0.3409(10)</i>	<i>0.4351(10)</i>	<i>0.4646(10)</i>	<i>0.4990(9)</i>	<i>0.6322(2)</i>

**AUC Insurance mcar** In grassetto i risultati migliori.

AUC ensemble learning Insurance mar						
Metodo	p	100	200	500	1000	5000
k10	1	<b>0.4731(10)</b>	0.5291(10)	<b>0.6483(10)</b>	0.6412(10)	0.7196(10)
	2	<b>0.4272(10)</b>	<b>0.5319(10)</b>	0.5882(10)	<b>0.6448(10)</b>	0.6635(10)
	3	0.4142(10)	<b>0.4845(10)</b>	0.5827(10)	<b>0.5983(10)</b>	0.6265(10)
SEM+T	1	0.4647(10)	0.5011(10)	0.5323(10)	0.6552(10)	0.6908(10)
	2	0.4226(10)	0.4549(10)	0.5125(10)	0.5340(10)	0.6208(10)
	3	<b>0.4270(10)</b>	0.4685(10)	0.5184(10)	0.5105(10)	0.5940(10)
k10 + SEM+T	1	0.4572(10)	<b>0.5436(10)</b>	0.6191(10)▽	<b>0.6757(10)</b>	<b>0.8111(10)</b>
	2	0.4068(10)	0.5049(10)	<b>0.5996(10)</b>	0.6414(10)	<b>0.7341(10)</b> ▽
	3	0.4247(10)	0.4566(10)	<b>0.5833(10)</b>	0.5957(10)	<b>0.6629(10)</b>
SEM+RT	1	<i>0.4372(10)</i>	<i>0.4708(10)</i>	<i>0.4928(10)</i>	<i>0.5678(10)</i>	<i>0.5925(10)</i>
	2	<i>0.3577(10)</i>	<i>0.4438(10)</i>	<i>0.4751(10)</i>	<i>0.5244(10)</i>	<i>0.5722(10)</i>
	3	<i>0.4078(10)</i>	<i>0.3907(10)</i>	<i>0.5162(10)</i>	<i>0.4493(10)</i>	<i>0.5447(10)</i>
k10 + SEM+RT	1	<i>0.4602(10)</i>	<i>0.5510(10)</i>	<i>0.5923(10)</i>	<i>0.6533(10)</i>	<i>0.7524(10)</i>
	2	<i>0.4120(10)</i>	<i>0.4895(10)</i>	<i>0.5414(10)</i>	<i>0.6365(10)</i>	<i>0.6833(10)</i>
	3	<i>0.4245(10)</i>	<i>0.4461(10)</i>	<i>0.5840(10)</i>	<i>0.5571(10)</i>	<i>0.6203(10)</i>

AUC Insurance mar In grassetto i risultati migliori.

AUC ensemble learning Insurance nmar						
Metodo	P	100	200	500	1000	5000
k10	1	<b>0.4552(10)</b>	<b>0.5371(10)</b>	0.5466(10)	0.6414(10)	0.7518(10)
	2	<b>0.4206(10)</b>	<b>0.4591(10)</b>	0.5189(10)	0.5925(10)	0.7094(10)
	3	0.4082(10)	<b>0.4642(10)</b>	<b>0.5145(10)</b>	0.5662(10)	0.6463(10)
SEM+T	1	0.3980(10)	0.4701(10)	0.5308(10)	0.5933(10)†	0.7269(10)†
	2	0.4016(10)	0.4071(10)	0.5001(10)	0.5883(10)	0.6015(8)
	3	<b>0.4353(10)</b>	0.4280(10)†	0.4811(10)	0.5102(10)	0.6283(10)†
k10 + SEM+T	1	0.4352(10)	0.5077(10)	<b>0.5598(10)</b>	<b>0.6585(10)</b> ▽	<b>0.8134(10)</b> ▽
	2	0.4056(10)	0.4407(10)	<b>0.5581(10)</b>	<b>0.6379(10)</b>	<b>0.7184(8)</b>
	3	0.4295(10)▽	0.4331(10)	0.5014(10)	<b>0.5869(10)</b>	<b>0.6911(10)</b> ▽
SEM+RT	1	<i>0.4239(10)</i>	<i>0.4626(10)</i>	<i>0.5170(10)</i>	<i>0.5456(10)</i>	<i>0.6171(10)</i>
	2	<i>0.3760(10)</i>	<i>0.4100(10)</i>	<i>0.5076(10)</i>	<i>0.5116(10)</i>	<i>0.5555(10)</i>
	3	<i>0.3755(10)</i>	<i>0.4009(10)</i>	<i>0.4858(10)</i>	<i>0.4993(10)</i>	<i>0.5193(10)</i>
k10 + SEM+RT	1	<i>0.4374(10)</i>	<i>0.5229(10)</i>	<i>0.5557(10)</i>	<i>0.6448(10)</i>	<i>0.7115(10)</i>
	2	<i>0.3903(10)</i>	<i>0.4456(10)</i>	<i>0.5453(10)</i>	<i>0.6014(10)</i>	<i>0.6661(10)</i>
	3	<i>0.3767(10)</i>	<i>0.4136(10)</i>	<i>0.5154(10)</i>	<i>0.5735(10)</i>	<i>0.6416(10)</i>

**AUC Insurance nmar** In grassetto i risultati migliori.

AUC ensemble learning Alarm1 mcar

Metodo	p	100	200	500	1000	5000
k10	1	0.4982(10)	0.5094(10)	0.6600(10)	0.7007(10)	0.7824(10)
	2	0.4032(10)	0.5057(10)	0.5870(10)	0.6298(10)	<b>0.7217(10)</b>
	3	0.3072(10)	0.3850(10)	0.4988(10)	0.5518(10)	0.6226(10)
SEM+T	1	<b>0.6905(10)*†</b>	0.6968(10)†	0.8302(10)†	0.8248(10)†	<b>0.9027(10)†</b>
	2	0.4687(10)†	0.5751(10)†	0.6784(10)†	0.7133(10)†	0.5051(8)
	3	<b>0.3714(10)†</b>	<b>0.3901(10)</b>	0.5401(10)†	0.6062(9)†	<b>0.7120(5)</b>
k10 + SEM+T	1	0.6682(10)∇	<b>0.7030(10)∇</b>	<b>0.8473(10)∇</b>	<b>0.8453(10)∇</b>	0.9013(10)∇
	2	<b>0.4787(10)∇</b>	<b>0.5804(10)∇</b>	<b>0.7201(10)∇</b>	<b>0.7815(10)∇</b>	0.6550(8)
	3	0.3400(10)∇	0.3693(10)	<b>0.5476(10)∇</b>	<b>0.6335(9)∇</b>	0.6646(5)
SEM+RT	1	0.5103(10)	0.5279(10)	0.5725(10)	0.6859(10)	0.6340(9)
	2	0.4367(10)	0.4324(10)	0.5797(10)	0.5288(9)	0.4893(5)
	3	0.2516(10)	0.3801(10)	0.3851(10)	0.4547(7)	-
k10 + SEM+RT	1	0.5811(10)	0.5575(10)	0.6943(10)	0.7693(10)	0.7743(9)
	2	0.4440(10)	0.5042(10)	0.6329(10)	0.6222(9)	0.6843(5)
	3	0.2708(10)	0.3767(10)	0.4728(10)	0.5554(7)	-

**AUC Alarm1 mcar** In grassetto i risultati migliori.

AUC ensemble learning Alarm1 mar						
Metodo	p	100	200	500	1000	5000
k10	1	0.4417(10)	0.5808(10)	0.6741(10)	0.6611(10)	0.7758(10)
	2	0.4505(10)	0.4653(10)	0.6324(10)	0.6539(10)	0.7744(10)
	3	0.4082(10)	0.4693(10)	0.6040(10)	0.6112(10)	0.6554(10)
SEM+T	1	<b>0.6117(10)*†</b>	<b>0.7154(10)†</b>	0.8038(10)†	0.7915(10)†	<b>0.9373(10)†</b>
	2	<b>0.5553(10)†</b>	<b>0.6380(9)†</b>	0.6599(10)†	0.6463(10)†	0.8437(10)†
	3	0.4908(10)†	<b>0.5958(9)†</b>	0.6415(10)†	0.7049(10)†	0.5955(10)
k10 + SEM+T	1	0.5490(10)▽	0.6986(10)▽	<b>0.8261(10)▽</b>	<b>0.8162(10)▽</b>	0.9285(10)▽
	2	0.5528(10)▽	0.6349(9)▽	<b>0.7197(10)▽</b>	<b>0.7388(10)</b>	<b>0.9036(10)*▽</b>
	3	<b>0.5141(10)▽</b>	0.5710(9)▽	<b>0.6973(10)▽</b>	<b>0.7655(10)▽</b>	<b>0.7334(10)</b>
SEM+RT	1	0.4688(10)	0.4234(10)	0.5596(10)	0.5834(10)	0.6370(10)
	2	0.4215(10)	0.3899(9)	0.5224(10)	0.5247(10)	0.5226(10)
	3	0.3276(10)	0.3711(10)	0.5213(10)	0.5139(10)	0.5182(10)
k10 + SEM+RT	1	0.4826(10)	0.5685(10)	0.6960(10)	0.6902(10)	0.8103(10)
	2	0.4652(10)	0.4819(9)	0.6159(10)	0.6813(10)	0.7163(10)
	3	0.4041(10)	0.4183(10)	0.6107(10)	0.6144(10)	0.6605(10)

**AUC Alarm1 mar** In grassetto i risultati migliori.

AUC ensemble learning Alarm1 nmar						
Metodo	p	100	200	500	1000	5000
k10	1	0.4635(10)	0.5625(10)	0.6260(10)	0.7312(10)	0.7630(10)
	2	0.4512(10)	0.5236(10)	0.6280(10)	0.6448(10)	0.7163(10)
	3	0.3808(10)	0.4579(10)	0.5776(10)	0.6207(10)	0.6630(10)
SEM+T	1	0.5927(10)†	0.6572(10)†	0.7506(10)†	0.8360(10)†	0.8686(10)†
	2	<b>0.5024(10)†</b>	<b>0.6405(10)†</b>	0.7236(10)†	0.7399(10)†	0.7168(10)†
	3	<b>0.4646(10)†</b>	<b>0.5941(10)†</b>	0.6195(10)†	0.6475(10)†	0.6950(10)†
k10 + SEM+T	1	<b>0.6062(10)▽</b>	<b>0.6804(10)▽</b>	<b>0.7773(10)▽</b>	<b>0.8695(10)*▽</b>	<b>0.9101(10)*▽</b>
	2	0.5024(10)▽	0.6268(10)▽	<b>0.7316(10)▽</b>	<b>0.7909(10)*▽</b>	<b>0.8243(10)▽</b>
	3	0.4581(10)▽	0.5803(10)▽	<b>0.6494(10)▽</b>	<b>0.7105(10)▽</b>	<b>0.7480(10)▽</b>
SEM+RT	1	0.4688(10)	0.5035(10)	0.5309(10)	0.5709(10)	0.5865(10)
	2	0.3625(10)	0.4643(10)	0.4455(10)	0.5613(10)	0.5295(10)
	3	0.3688(10)	0.4014(10)	0.4696(10)	0.4606(10)	0.5068(10)
k10 + SEM+RT	1	0.4873(10)	0.5734(10)	0.6505(10)	0.7406(10)	0.7563(10)
	2	0.4342(10)	0.5320(10)	0.5948(10)	0.6566(10)	0.7076(10)
	3	0.3870(10)	0.4341(10)	0.5820(10)	0.5945(10)	0.6243(10)

AUC Alarm1 nmar In grassetto i risultati migliori.

AUC ensemble learning HailFinder mcar						
Metodo	p	100	200	500	1000	5000
k10	1	<b>0.0846(10)</b>	<b>0.0913(10)</b>	0.0998(10)	0.0891(10)	0.1237(10)
	2	0.0846(10)	0.0922(10)	0.0879(10)	<b>0.1064(10)</b>	0.1092(10)
	3	0.0698(10)	0.0804(10)	0.0867(10)	0.0811(10)	<b>0.1111(10)</b>
SEM+T	1	0.0827(10)	0.0876(10)	<b>0.1081(10)</b>	<b>0.1190(10)*†</b>	<b>0.1488(8)*</b>
	2	<b>0.0880(10)†</b>	<b>0.0926(10)</b>	<b>0.1122(10)</b>	0.1013(8)	<b>0.1227(2)</b>
	3	<b>0.1013(10)*†</b>	<b>0.0936(10)†</b>	<b>0.1065(10)†</b>	<b>0.1075(10)*†</b>	-
k10 + SEM+T	1	0.0777(10)∇	0.0810(10)	0.0901(10)	0.0885(10)∇	0.1187(8)
	2	0.0750(10)∇	0.0835(10)	0.0831(10)	0.0852(8)	0.0990(2)
	3	0.0669(10)∇	0.0752(10)∇	0.0835(10)∇	0.0799(10)∇	-
SEM+RT	1	<i>0.0679(10)</i>	<i>0.0800(10)</i>	<i>0.1044(10)</i>	<i>0.0957(10)</i>	<i>0.1305(8)</i>
	2	<i>0.0536(10)</i>	<i>0.0884(10)</i>	<i>0.0884(10)</i>	<i>0.1156(8)</i>	-
	3	<i>0.0132(10)</i>	<i>0.0536(10)</i>	<i>0.0731(9)</i>	<i>0.0957(8)</i>	-
k10 + SEM+RT	1	<i>0.0711(10)</i>	<i>0.0749(10)</i>	<i>0.0858(10)</i>	<i>0.0859(10)</i>	<i>0.1093(8)</i>
	2	<i>0.0615(10)</i>	<i>0.0777(10)</i>	<i>0.0799(10)</i>	<i>0.0936(8)</i>	-
	3	<i>0.0445(10)</i>	<i>0.0578(10)</i>	<i>0.0733(9)</i>	<i>0.0770(8)</i>	-

**AUC HailFinder mcar** In grassetto i risultati migliori.

AUC ensemble learning HailFinder mar

Metodo	P	100	200	500	1000	5000
k10	1	<b>0.0842(10)</b>	0.0884(10)	0.0841(10)	0.0930(10)	0.1008(10)
	2	0.0822(10)	0.0874(10)	0.0936(10)	0.0938(10)	0.1070(10)
	3	0.0693(10)	0.0776(10)	0.0819(10)	0.0906(10)	0.1099(10)
SEM+T	1	0.0803(10)†	<b>0.0884(10)†</b>	<b>0.1000(10)</b>	<b>0.1098(10)*†</b>	<b>0.1300(10)*</b>
	2	<b>0.0855(10)†</b>	<b>0.0906(10)†</b>	<b>0.0978(10)</b>	<b>0.1134(9)†</b>	<b>0.1217(9)</b>
	3	<b>0.0843(10)†</b>	<b>0.0916(10)*†</b>	<b>0.1080(10)*†</b>	<b>0.1155(10)*†</b>	<b>0.1322(7)*†</b>
k10 + SEM+T	1	0.0749(10)∇	0.0801(10)∇	0.0777(10)	0.0849(10)∇	0.1008(10)
	2	0.0730(10)∇	0.0785(10)∇	0.0856(10)	0.0978(9)	0.0959(9)
	3	0.0660(10)∇	0.0769(10)∇	0.0813(10)∇	0.0882(10)∇	0.1057(7)∇
SEM+RT	1	<i>0.0664(10)</i>	<i>0.0709(10)</i>	<i>0.0848(10)</i>	<i>0.0980(10)</i>	<i>0.1327(10)</i>
	2	<i>0.0500(10)</i>	<i>0.0759(10)</i>	<i>0.0838(10)</i>	<i>0.0850(10)</i>	<i>0.1172(7)</i>
	3	<i>0.0520(10)</i>	<i>0.0638(10)</i>	<i>0.0757(10)</i>	<i>0.0867(10)</i>	<i>0.1080(8)</i>
k10 + SEM+RT	1	<i>0.0699(10)</i>	<i>0.0683(10)</i>	<i>0.0777(10)</i>	<i>0.0827(10)</i>	<i>0.1005(10)</i>
	2	<i>0.0641(10)</i>	<i>0.0727(10)</i>	<i>0.0774(10)</i>	<i>0.0831(10)</i>	<i>0.0948(7)</i>
	3	<i>0.0531(10)</i>	<i>0.0623(10)</i>	<i>0.0712(10)</i>	<i>0.0779(10)</i>	<i>0.0949(8)</i>

**AUC HailFinder mar** In grassetto i risultati migliori.

AUC ensemble learning HailFinder nmar						
Metodo	p	100	200	500	1000	5000
k10	1	<b>0.0864(10)</b>	<b>0.0909(10)</b>	0.0871(10)	0.0943(10)	0.0939(10)
	2	0.0768(10)	<b>0.0890(10)</b>	0.0874(10)	0.0846(10)	0.1075(10)
	3	0.0740(10)	0.0805(10)	0.0850(10)	0.0791(10)	0.0961(10)
SEM+T	1	0.0855(10)	0.0842(10)†	<b>0.0992(10)</b>	<b>0.1122(10)†</b>	<b>0.1443(7)*</b>
	2	<b>0.0812(10)†</b>	0.0880(10)†	<b>0.1002(10)†</b>	<b>0.1077(10)</b>	<b>0.1312(5)</b>
	3	<b>0.0805(10)†</b>	<b>0.0936(10)†</b>	<b>0.1051(10)†</b>	<b>0.1022(10)*†</b>	<b>0.1114(6)</b>
k10 + SEM+T	1	0.0807(10)	0.0817(10)∇	0.0850(10)	0.0845(10)∇	0.1024(7)
	2	0.0774(10)∇	0.0798(10)∇	0.0804(10)	0.0842(10)	0.1003(5)
	3	0.0664(10)∇	0.0772(10)∇	0.0825(10)∇	0.0750(10)∇	0.0865(6)
SEM+RT	1	<i>0.0684(10)</i>	<i>0.0690(10)</i>	<i>0.0927(10)</i>	<i>0.0947(10)</i>	<i>0.1274(9)</i>
	2	<i>0.0536(10)</i>	<i>0.0738(10)</i>	<i>0.0766(10)</i>	<i>0.0905(10)</i>	<i>0.1267(8)</i>
	3	<i>0.0544(10)</i>	<i>0.0654(10)</i>	<i>0.0874(10)</i>	<i>0.0830(10)</i>	<i>0.0945(7)</i>
k10 + SEM+RT	1	<i>0.0733(10)</i>	<i>0.0735(10)</i>	<i>0.0812(10)</i>	<i>0.0796(10)</i>	<i>0.1018(9)</i>
	2	<i>0.0650(10)</i>	<i>0.0742(10)</i>	<i>0.0748(10)</i>	<i>0.0768(10)</i>	<i>0.1010(8)</i>
	3	<i>0.0561(10)</i>	<i>0.0683(10)</i>	<i>0.0750(10)</i>	<i>0.0664(10)</i>	<i>0.0745(7)</i>

**AUC HailFinder nmar** In grassetto i risultati migliori.

AUC ensemble learning Hepar2 mcar

Metodo	p	100	200	500	1000	5000
k10	1	0.1420(10)	0.1992(10)	0.3465(10)	0.4189(10)	0.6120(10)
	2	0.1127(10)	0.1839(10)	0.2649(10)	0.3612(10)	0.5325(10)
	3	0.0995(10)	0.1381(10)	0.2331(10)	0.3028(10)	0.5330(10)
SEM+T	1	<b>0.3044(10)*</b>	<b>0.3661(10)*</b>	<b>0.4651(10)*</b>	<b>0.5126(10)*</b>	0.5803(10)
	2	<b>0.2893(10)*</b>	<b>0.3467(10)*</b>	<b>0.3827(10)*</b>	<b>0.4618(10)*</b>	<b>0.5738(10)</b>
	3	<b>0.3356(10)*</b>	<b>0.3290(10)*<sup>†</sup></b>	<b>0.3394(10)*</b>	<b>0.4217(10)*</b>	0.5457(10)
k10 + SEM+T	1	0.1483(10)	0.2285(10)	0.3803(10)	0.4673(10)	<b>0.6232(10)</b>
	2	0.1276(10)	0.2103(10) $\nabla$	0.2916(10)	0.3990(10)	0.5689(10)
	3	0.1163(10)	0.1602(10) $\nabla$	0.2392(10) $\nabla$	0.3366(10) $\nabla$	<b>0.5686(10)</b>
SEM+RT	1	<i>0.3099(10)</i>	<i>0.3529(10)</i>	<i>0.4406(10)</i>	<i>0.4829(10)</i>	<i>0.6343(10)<sup>†</sup></i>
	2	<i>0.2874(10)</i>	<i>0.3297(10)</i>	<i>0.3724(10)</i>	<i>0.4599(10)</i>	<i>0.6359(10)<sup>†</sup></i>
	3	<i>0.2745(10)</i>	<i>0.2911(10)</i>	<i>0.3169(10)</i>	<i>0.4006(10)</i>	<i>0.5739(10)</i>
k10 + SEM+RT	1	<i>0.1556(10)</i>	<i>0.2287(10)</i>	<i>0.3682(10)</i>	<i>0.4663(10)</i>	<i>0.6372(10)<math>\nabla</math></i>
	2	<i>0.1213(10)</i>	<i>0.2006(10)</i>	<i>0.2827(10)</i>	<i>0.4001(10)</i>	<i>0.6097(10)<math>\nabla</math></i>
	3	<i>0.1047(10)</i>	<i>0.1443(10)</i>	<i>0.2321(10)</i>	<i>0.3261(10)</i>	<i>0.5763(10)</i>

**AUC Hepar2 mcar** In grassetto i risultati migliori.

		AUC ensemble learning Hepar2 mar				
Metodo	p	100	200	500	1000	5000
k10	1	0.1274(10)	0.1757(10)	0.3115(10)	0.3870(10)	0.5597(10)
	2	0.1244(10)	0.1968(10)	0.2834(10)	0.3814(10)	0.4985(10)
	3	0.1128(10)	0.1604(10)	0.2911(10)	0.3326(10)	0.4638(10)
SEM+T	1	<b>0.2985(10)*</b>	<b>0.3320(10)*</b>	<b>0.4119(10)*</b>	<b>0.4676(10)*</b>	0.5381(10)
	2	<b>0.2930(10)*</b>	<b>0.3355(10)*</b>	<b>0.4111(10)*</b>	<b>0.4701(10)*</b>	0.5582(10)
	3	<b>0.2828(10)*</b>	<b>0.3408(10)*†</b>	<b>0.4074(10)*</b>	<b>0.4422(10)*</b>	<b>0.5657(6)</b>
k10 + SEM+T	1	0.1383(10)	0.2005(10)	0.3347(10)	0.4190(10)	<b>0.5882(10)</b>
	2	0.1447(10)	0.2112(10)	0.3127(10)	0.4289(10)	<b>0.5721(10)</b>
	3	0.1276(10)	0.1847(10)∇	0.3155(10)	0.3600(10)	0.5562(6)
SEM+RT	1	<i>0.2966(10)</i>	<i>0.3405(10)</i>	<i>0.4024(10)</i>	<i>0.4938(10)</i>	<i>0.6033(10)</i>
	2	<i>0.2648(10)</i>	<i>0.3405(10)</i>	<i>0.4283(10)</i>	<i>0.4776(10)</i>	<i>0.5940(10)</i>
	3	<i>0.2893(10)</i>	<i>0.3297(10)</i>	<i>0.3943(10)</i>	<i>0.4352(10)</i>	<i>0.5662(10)</i>
k10 + SEM+RT	1	<i>0.1373(10)</i>	<i>0.1993(10)</i>	<i>0.3367(10)</i>	<i>0.4376(10)</i>	<i>0.6055(10)</i>
	2	<i>0.1387(10)</i>	<i>0.2072(10)</i>	<i>0.3245(10)</i>	<i>0.4344(10)</i>	<i>0.5777(10)</i>
	3	<i>0.1267(10)</i>	<i>0.1767(10)</i>	<i>0.3180(10)</i>	<i>0.3684(10)</i>	<i>0.5402(10)</i>

**AUC Hepar2 mar** In grassetto i risultati migliori.

AUC ensemble learning Hepar2 nmar

Metodo	p	100	200	500	1000	5000
k10	1	0.1192(10)	0.1858(10)	0.3255(10)	0.4143(10)	0.5414(10)
	2	0.1240(10)	0.1833(10)	0.3062(10)	0.3806(10)	0.5382(9)
	3	0.1106(10)	0.1609(10)	0.2613(10)	0.3338(10)	0.5300(9)
SEM+T	1	<b>0.2985(10)*</b>	<b>0.3380(10)*</b>	<b>0.3847(10)*</b>	<b>0.4744(10)*</b>	0.5881(10)
	2	<b>0.3031(10)*</b>	<b>0.3348(10)*</b>	<b>0.4242(10)*</b>	<b>0.4851(10)*</b>	0.5673(10)
	3	<b>0.2893(10)*</b>	<b>0.3504(10)*†</b>	<b>0.3908(10)*</b>	<b>0.4526(10)*</b>	0.5671(10)
k10 + SEM+T	1	0.1291(10)	0.2130(10)	0.3451(10)	0.4428(10)	<b>0.5921(10)</b>
	2	0.1294(10)	0.2131(10)	0.3310(10)	0.4316(10)	<b>0.5850(9)</b>
	3	0.1137(10)	0.1869(10)∇	0.2988(10)	0.3807(10)	<b>0.5671(9)</b>
SEM+RT	1	<i>0.2985(10)</i>	<i>0.3447(10)</i>	<i>0.4581(10)</i>	<i>0.4777(10)</i>	<i>0.6296(10)</i>
	2	<i>0.2837(10)</i>	<i>0.3313(10)</i>	<i>0.4158(10)</i>	<i>0.4825(10)</i>	<i>0.5909(10)</i>
	3	<i>0.2893(10)</i>	<i>0.3247(10)</i>	<i>0.3865(10)</i>	<i>0.4960(10)</i>	<i>0.5915(10)</i>
k10 + SEM+RT	1	<i>0.1279(10)</i>	<i>0.2130(10)</i>	<i>0.3723(10)</i>	<i>0.4470(10)</i>	<i>0.6059(10)</i>
	2	<i>0.1294(10)</i>	<i>0.2080(10)</i>	<i>0.3366(10)</i>	<i>0.4149(10)</i>	<i>0.5974(9)</i>
	3	<i>0.1119(10)</i>	<i>0.1709(10)</i>	<i>0.2866(10)</i>	<i>0.4014(10)</i>	<i>0.5712(9)</i>

**AUC Hepar2 nmar** In grassetto i risultati migliori.

### 3.2.2 Grafici

In questa sezione si mostrano i grafici relativi alle tabella appena presentate. Una prima sezione sarà dedicata ai grafici di precision, recall e SHD, riportati su una stessa riga per ogni tipo e probabilità di missingness. Ogni grafico rappresenta una misura per tutti i metodi al variare della dimensione del dataset. I metodi indicati con linea puntata sono quelli che partono da rete vuota, quelli con linea tratto-punto sono inizializzati con rete MWST-EM. Si riporta anche in questo caso i risultati di MMHC con dataset completi come utile confronto e li si indica con linea continua e quadrato. I grafici sono disposti in modo da permettere il confronto tra reti diverse nelle medesime condizioni di missingness.

Considerazioni simili si applicano per i grafici relativi al tempo, nel qual caso, però, i grafici relativi alle varie reti sono condensati in una sola facciata per medesima condizione di incompletezza.

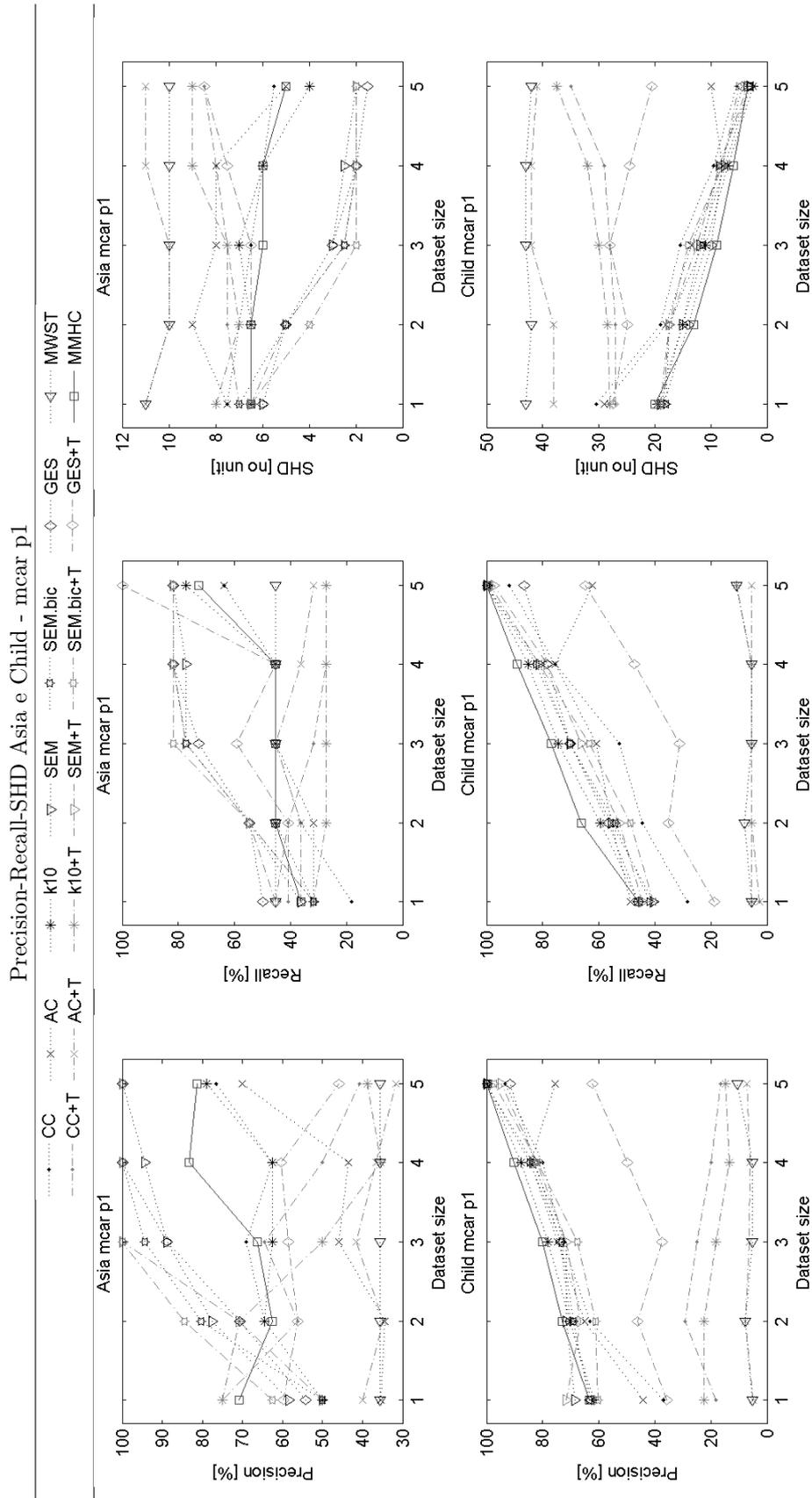
Si ricorda che per HailFinder ed Hepar2 il numero di metodi testato è minore rispetto alle altre quattro reti.

La terza sezione riporta i valori delle AUC ottenute con più metodi applicando l'approccio greedy precedentemente descritto alle quattro reti più piccole (Asia, Child, Insurance ed Alarm). Anche in questo caso, in una stessa facciata sono mostrate tutte le reti sotto le stesse condizioni di incompletezza.

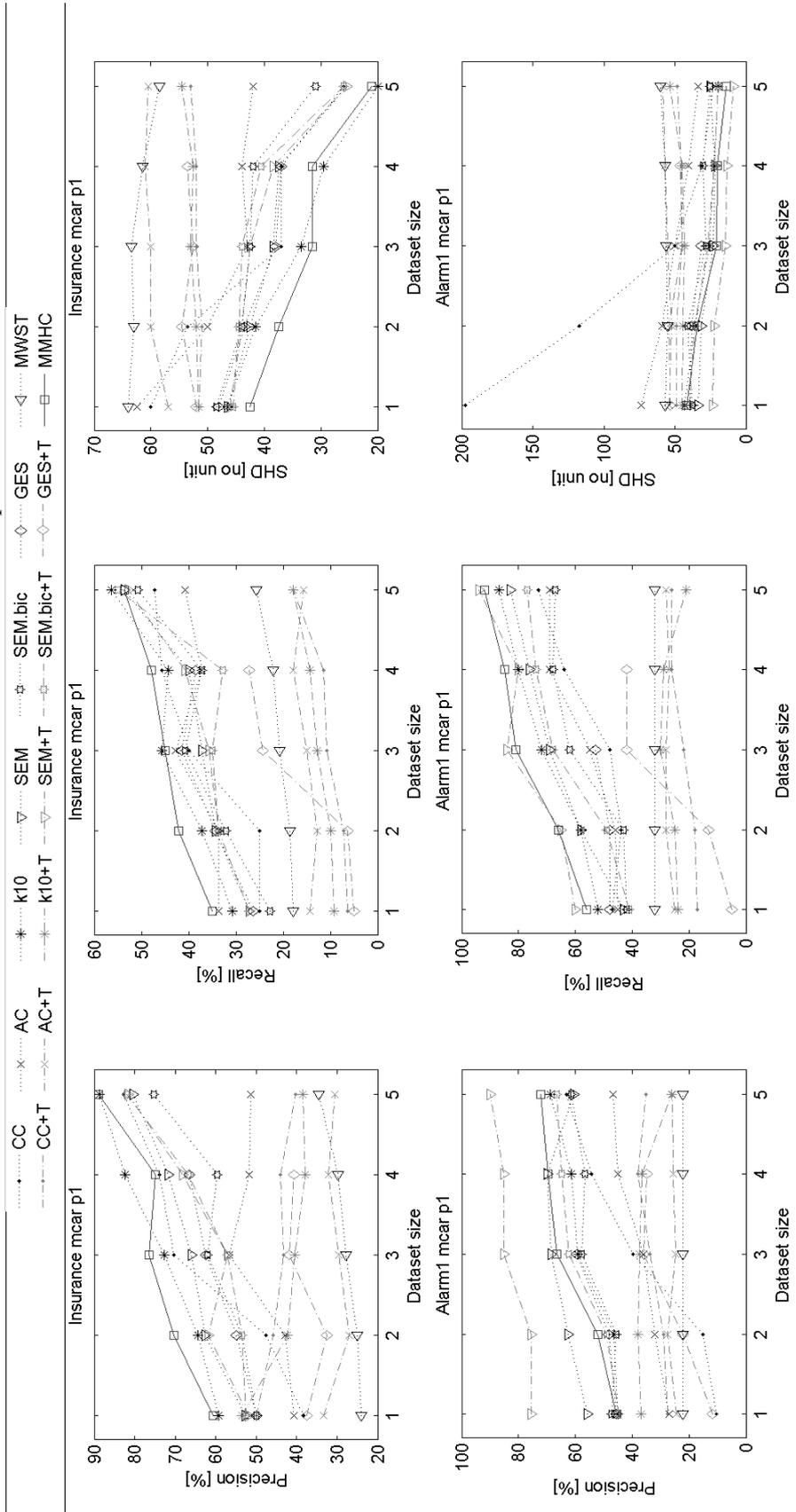
Infine, i grafici dell'ultima sezione mostrano per tutte le sei reti le AUC ottenute con i singoli metodi SEM+T e k-nn e come queste variano con la somma dei due metodi.

La taglia di dataset è riportata su scala da 1 a 5, dove le dimensioni dei dataset sono disposte in ordine crescente: 1 indica i dataset di taglia 100, 2 di taglia 200, 3 di taglia 500, 4 di taglia 1000 e 5 di taglia 5000.

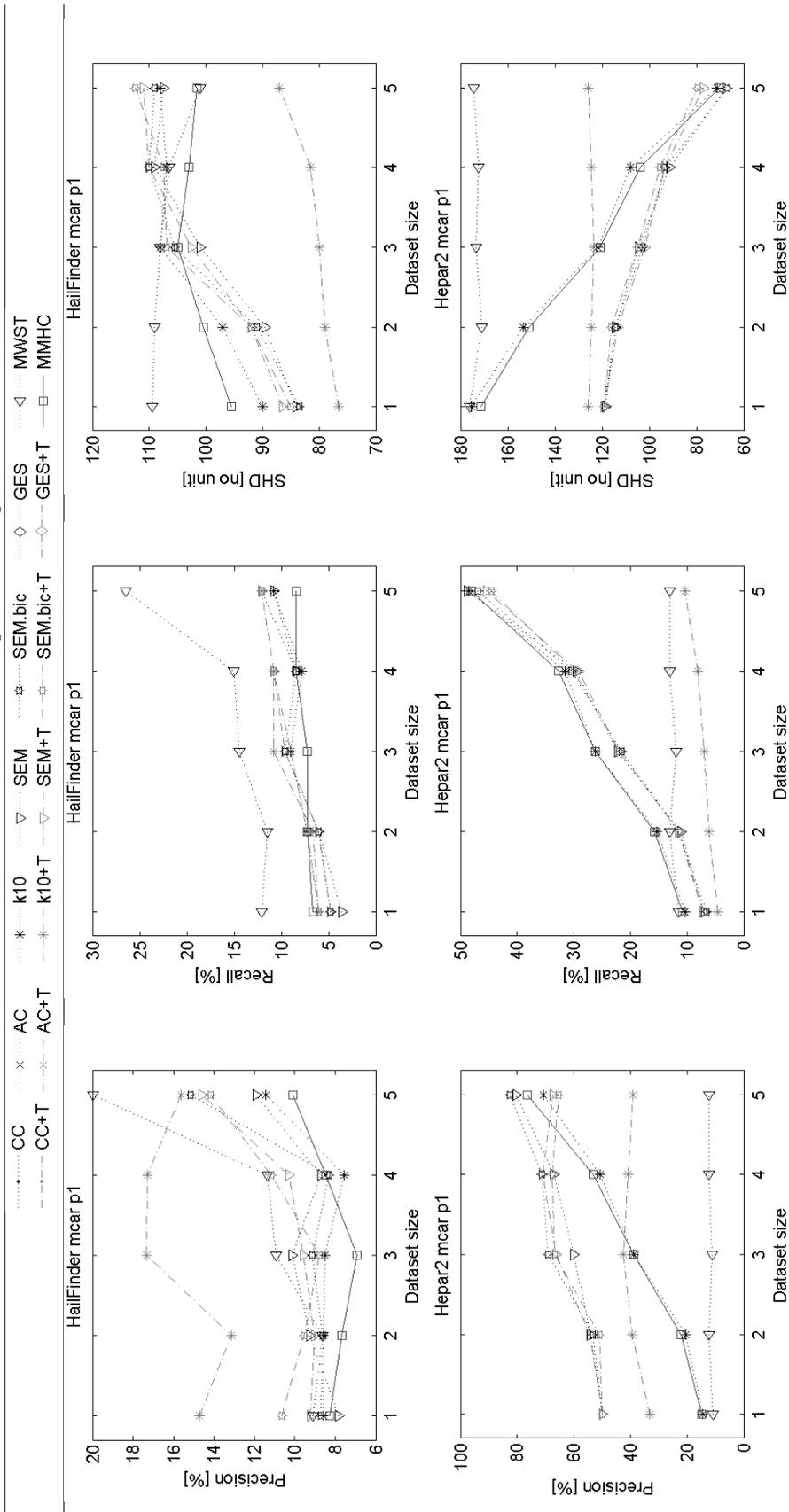
P-R-SHD



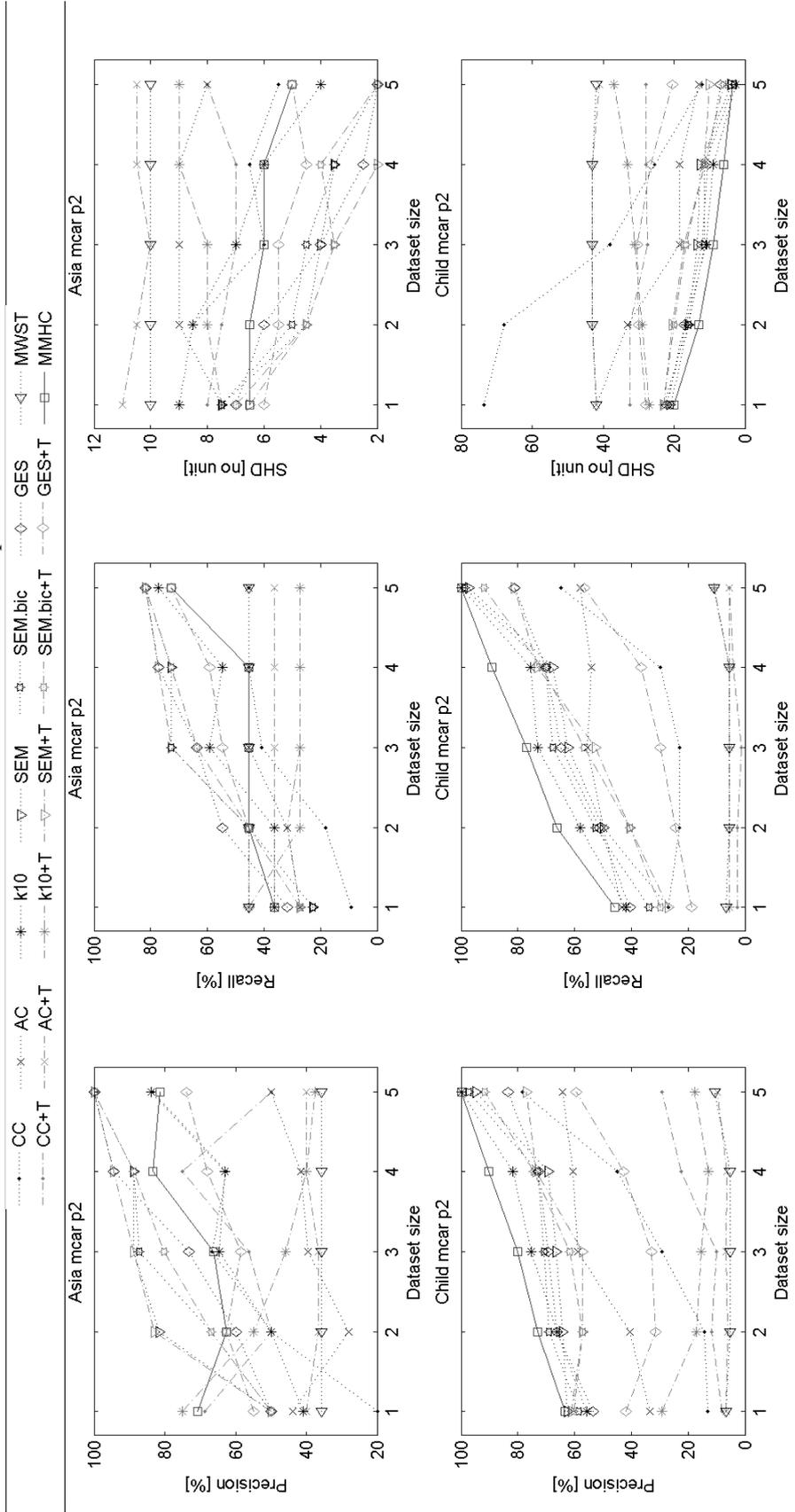
Precision-Recall-SHD Insurance e Alarm1 - mcar p1



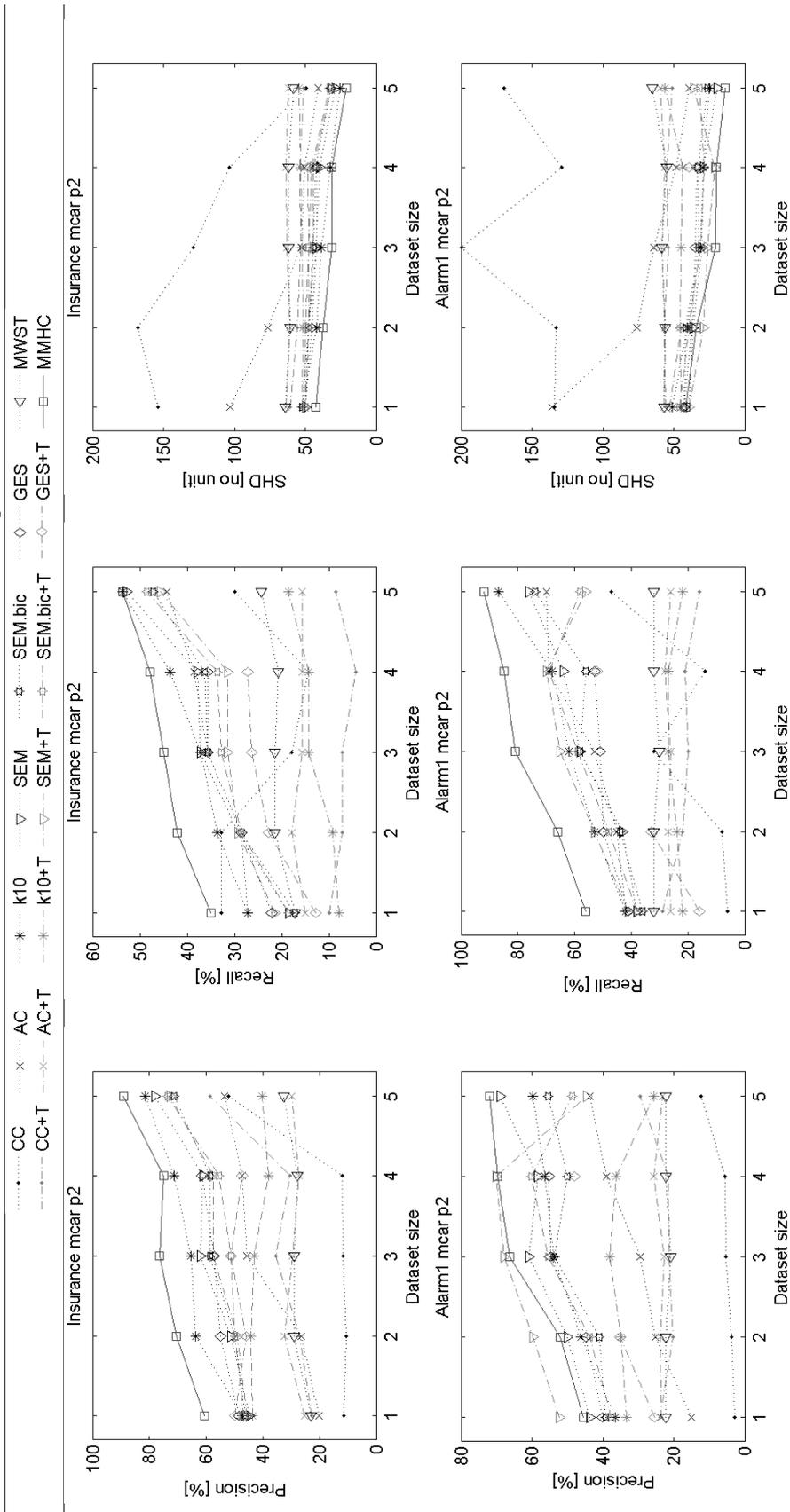
Precision-Recall-SHD HailFinder e Hepar2 - mcar p1



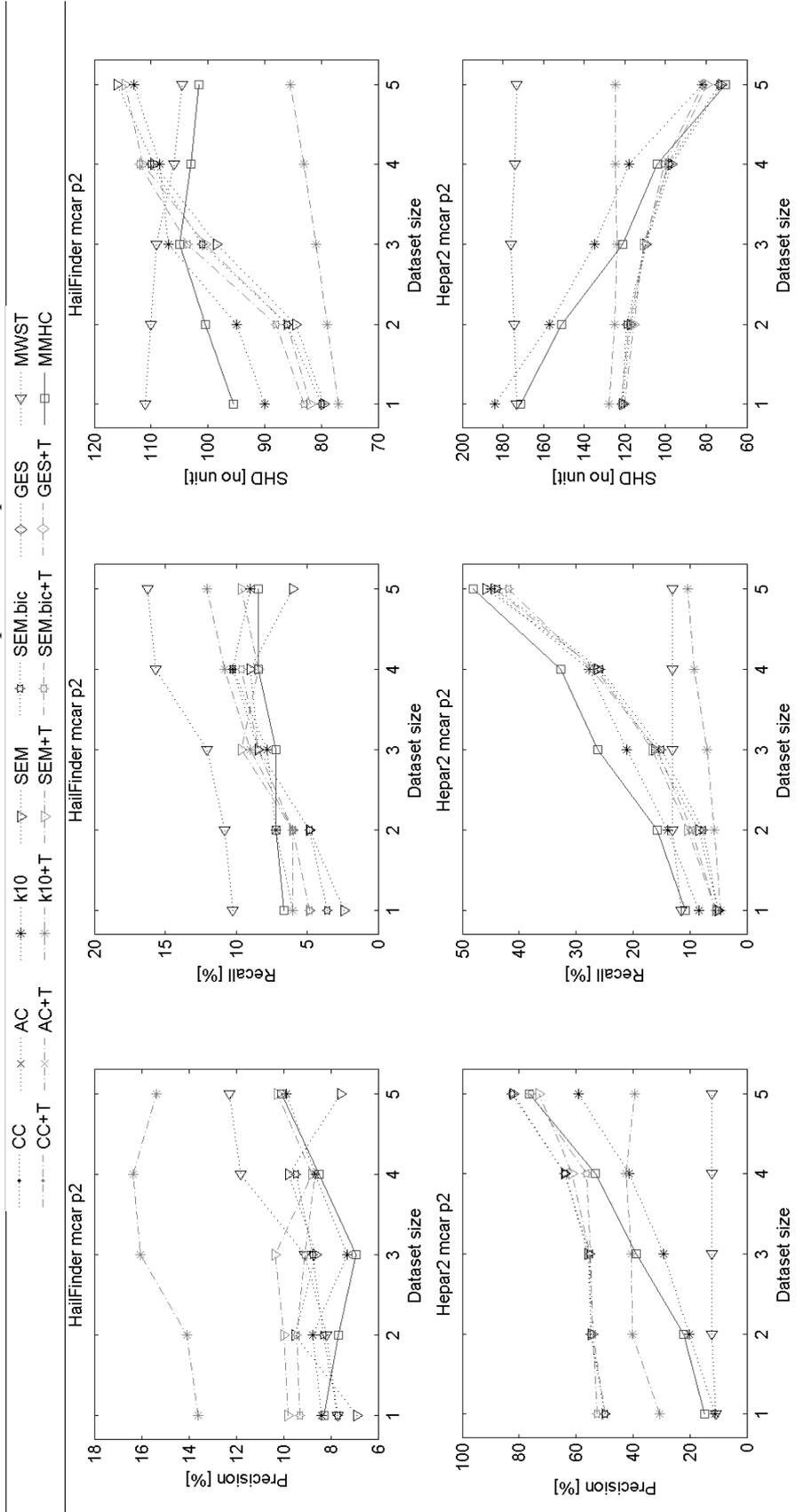
Precision-Recall-SHD Asia e Child - mcar p2



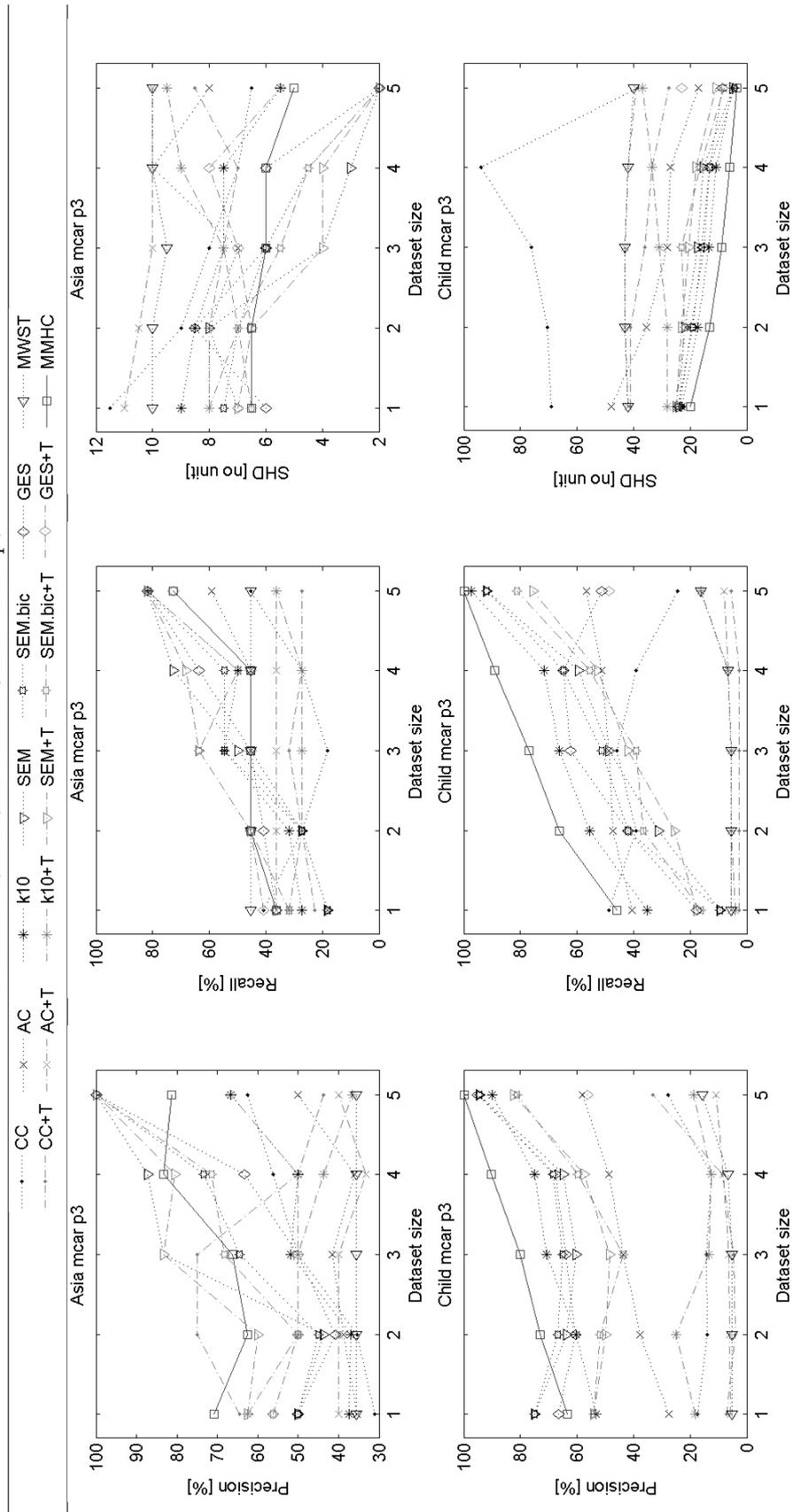
Precision-Recall-SHD Insurance e Alarm1 - mcar p2



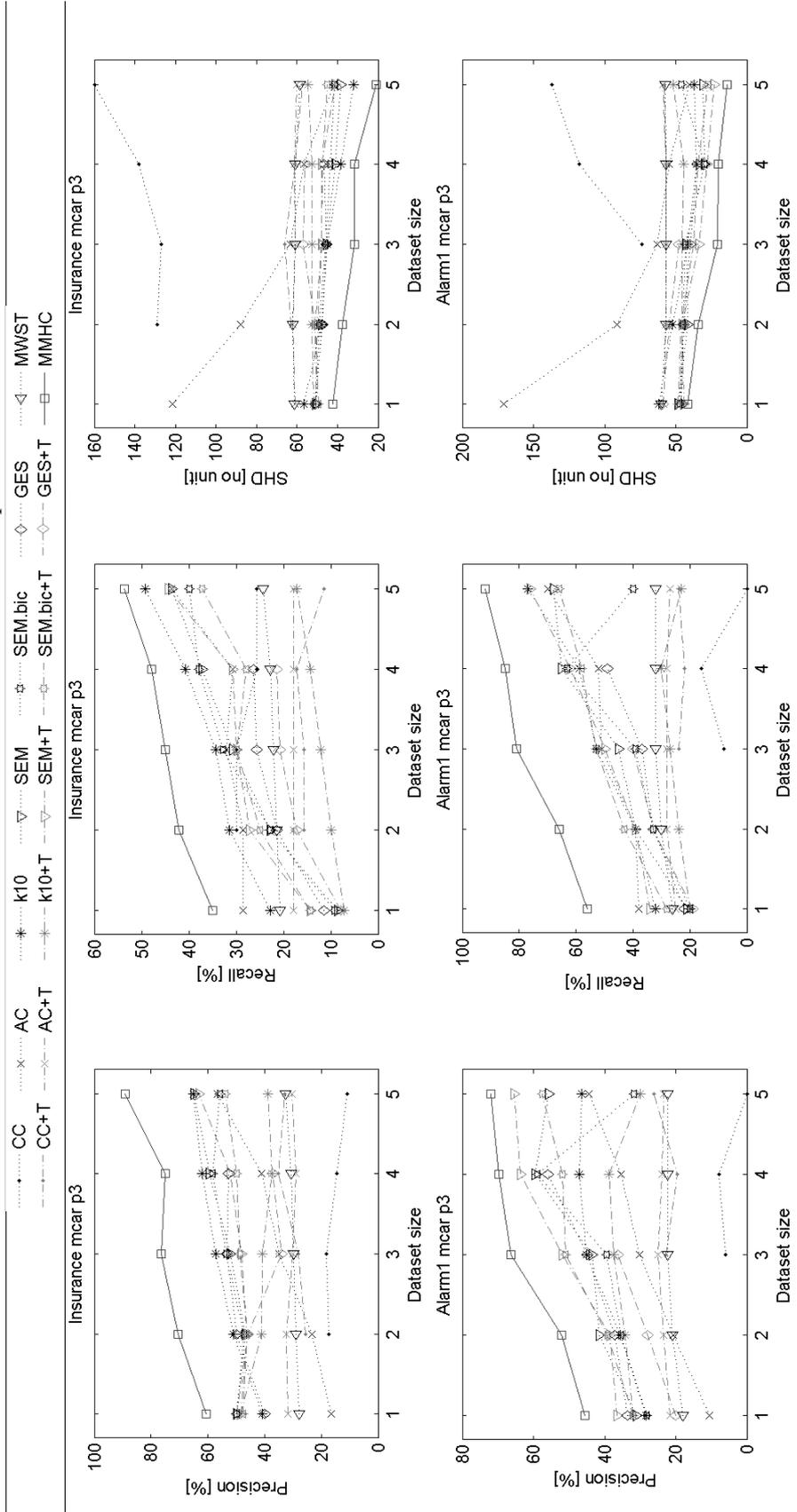
Precision-Recall-SHD HailFinder e Hepar2 - mcar p2



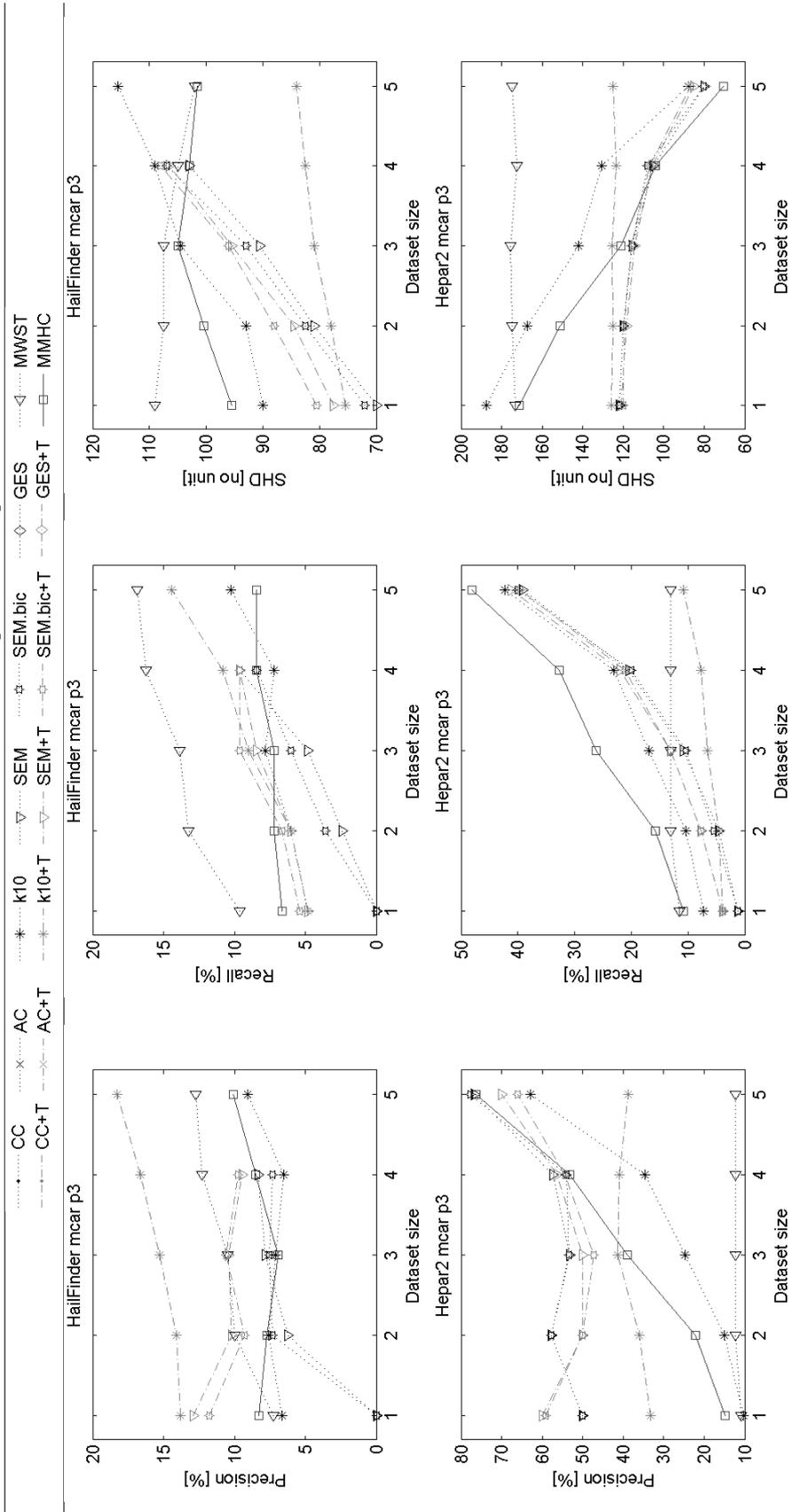
Precision-Recall-SHD Asia e Child - mcar p3



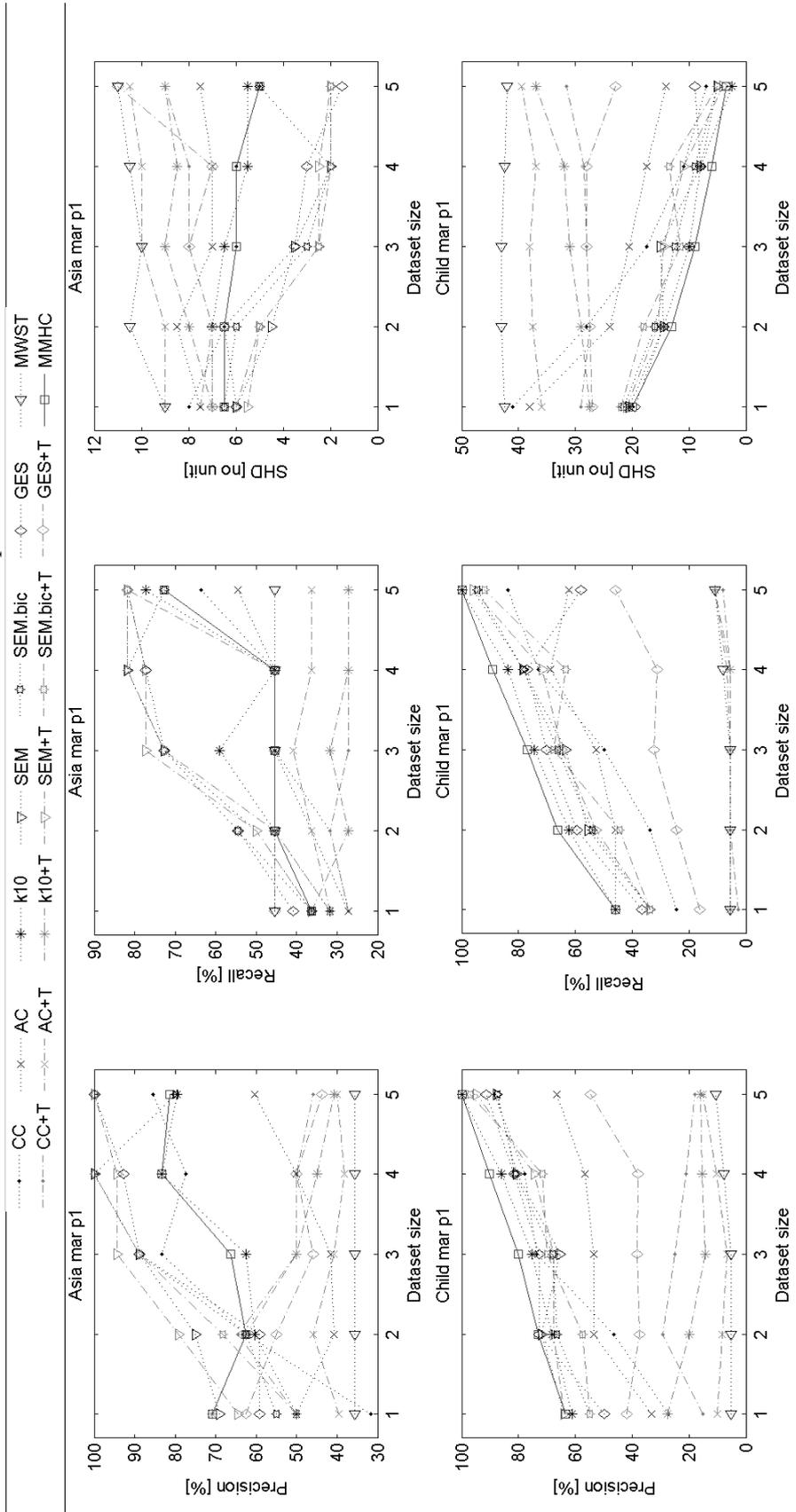
Precision-Recall-SHD Insurance e Alarm1 - mcar p3



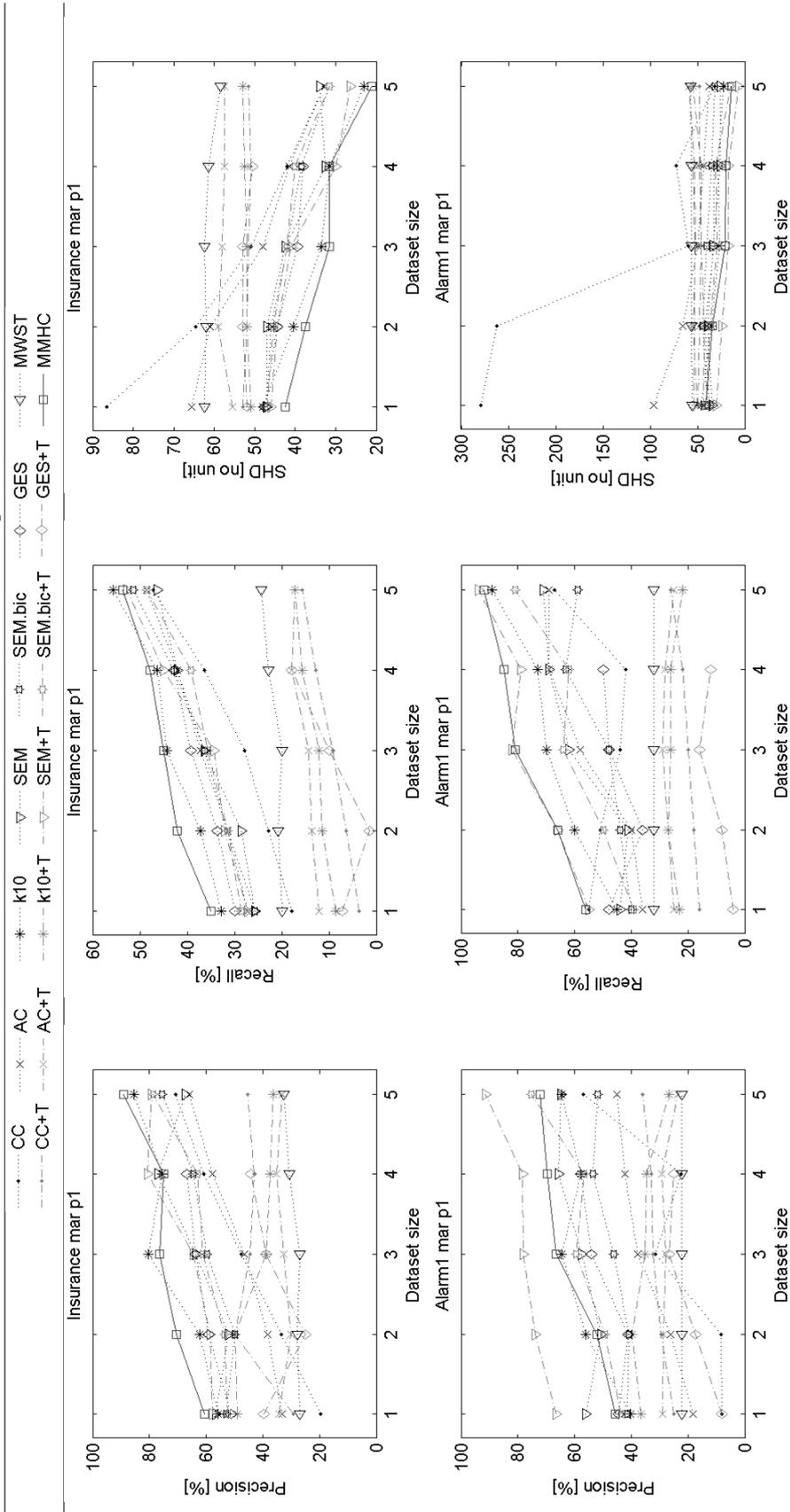
Precision-Recall-SHD HailFinder e Hepar2 - mcar p3



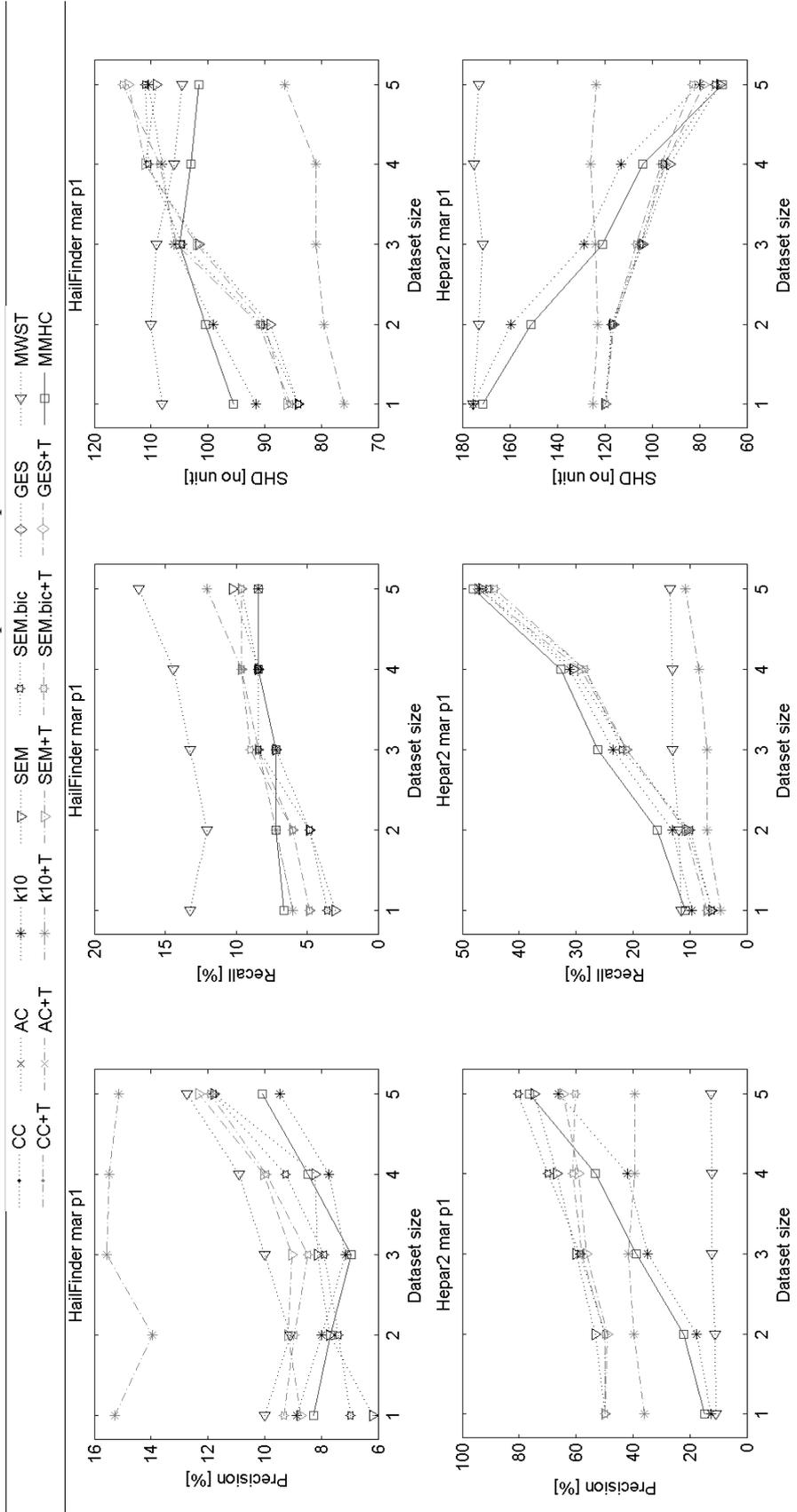
Precision-Recall-SHD Asia e Child - mar p1



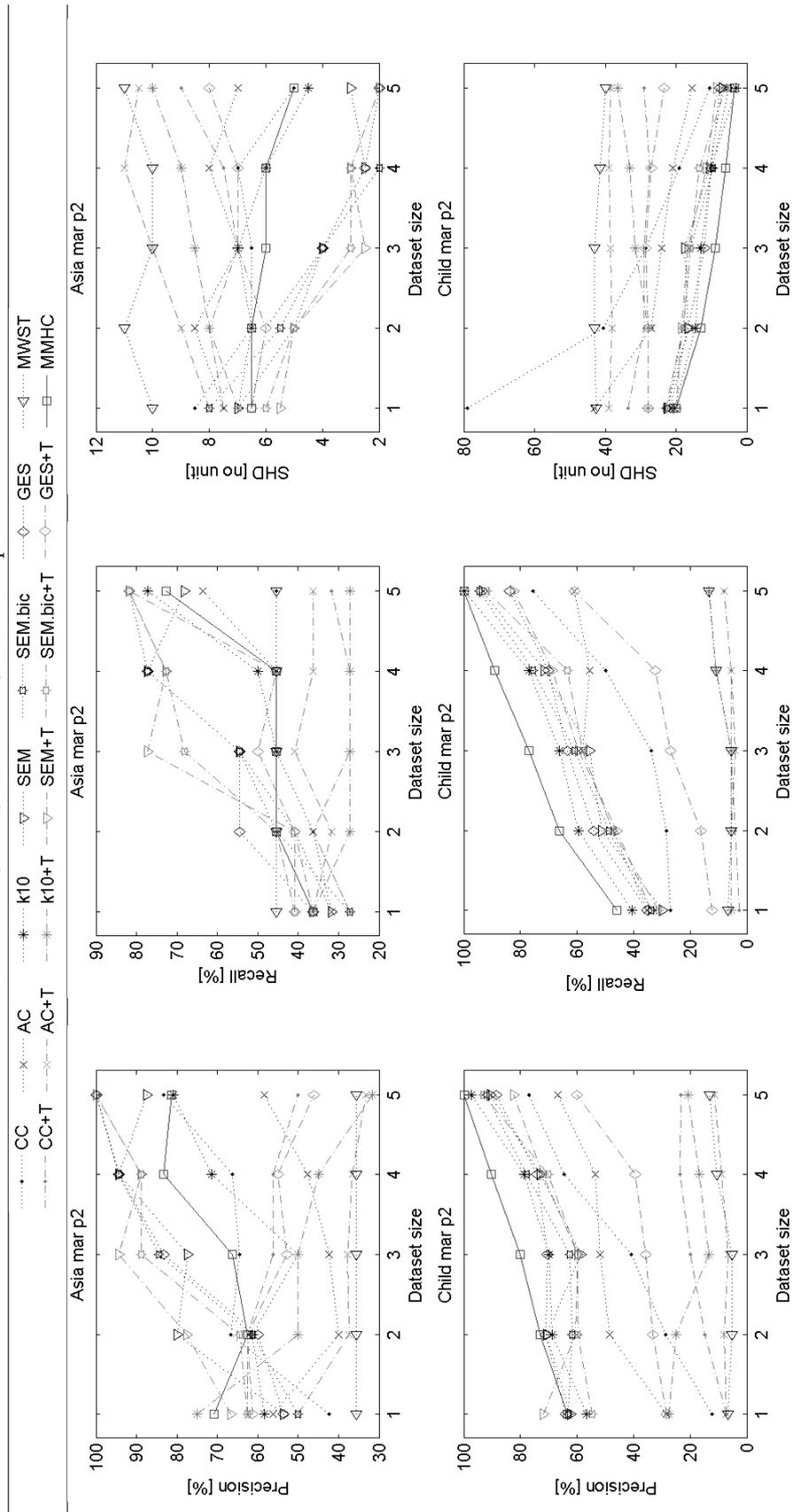
Precision-Recall-SHD Insurance e Alarm1 - mar p1



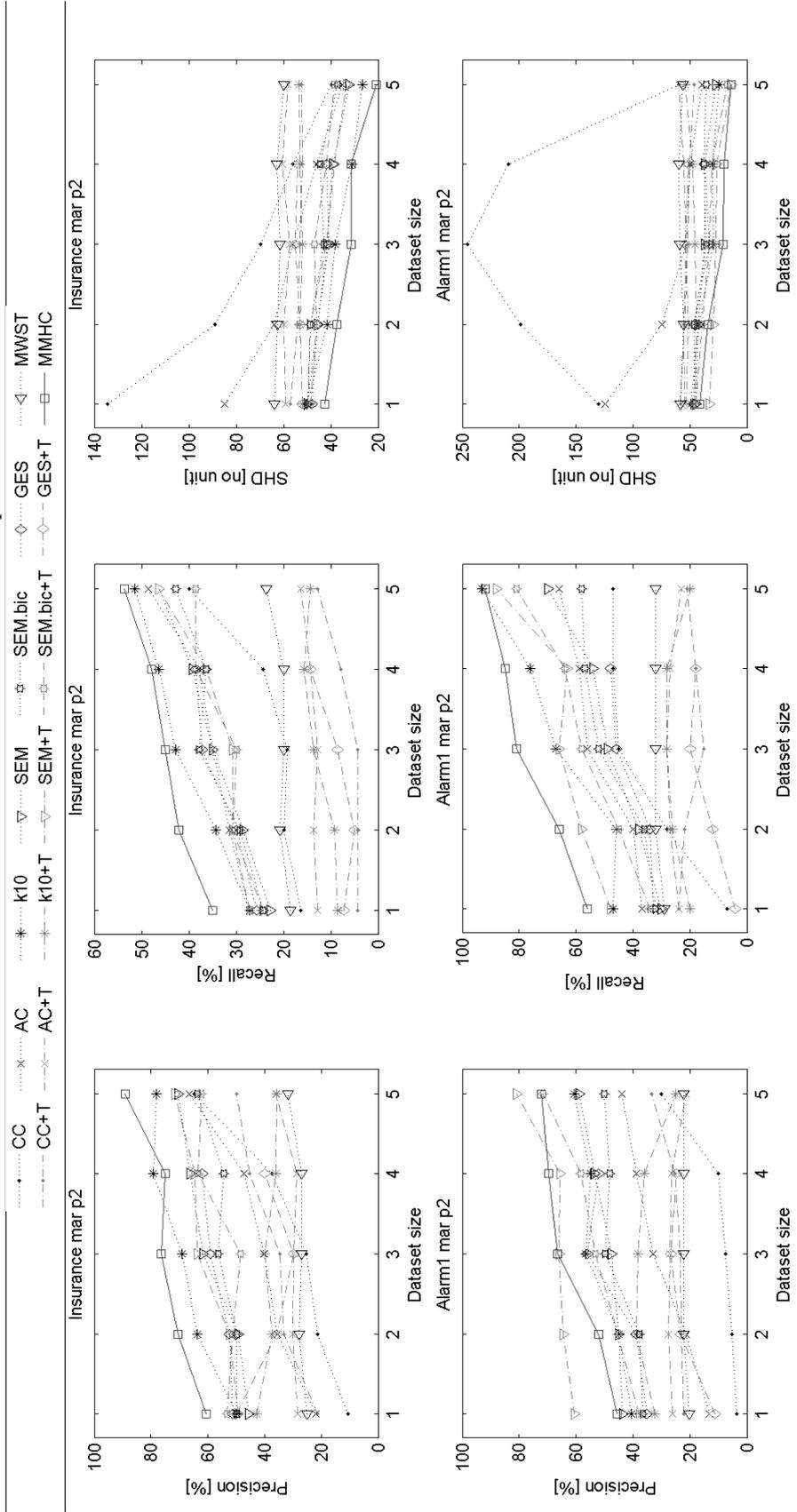
Precision-Recall-SHD HailFinder e Hepar2 - mar p1



Precision-Recall-SHD Asia e Child - mar p2

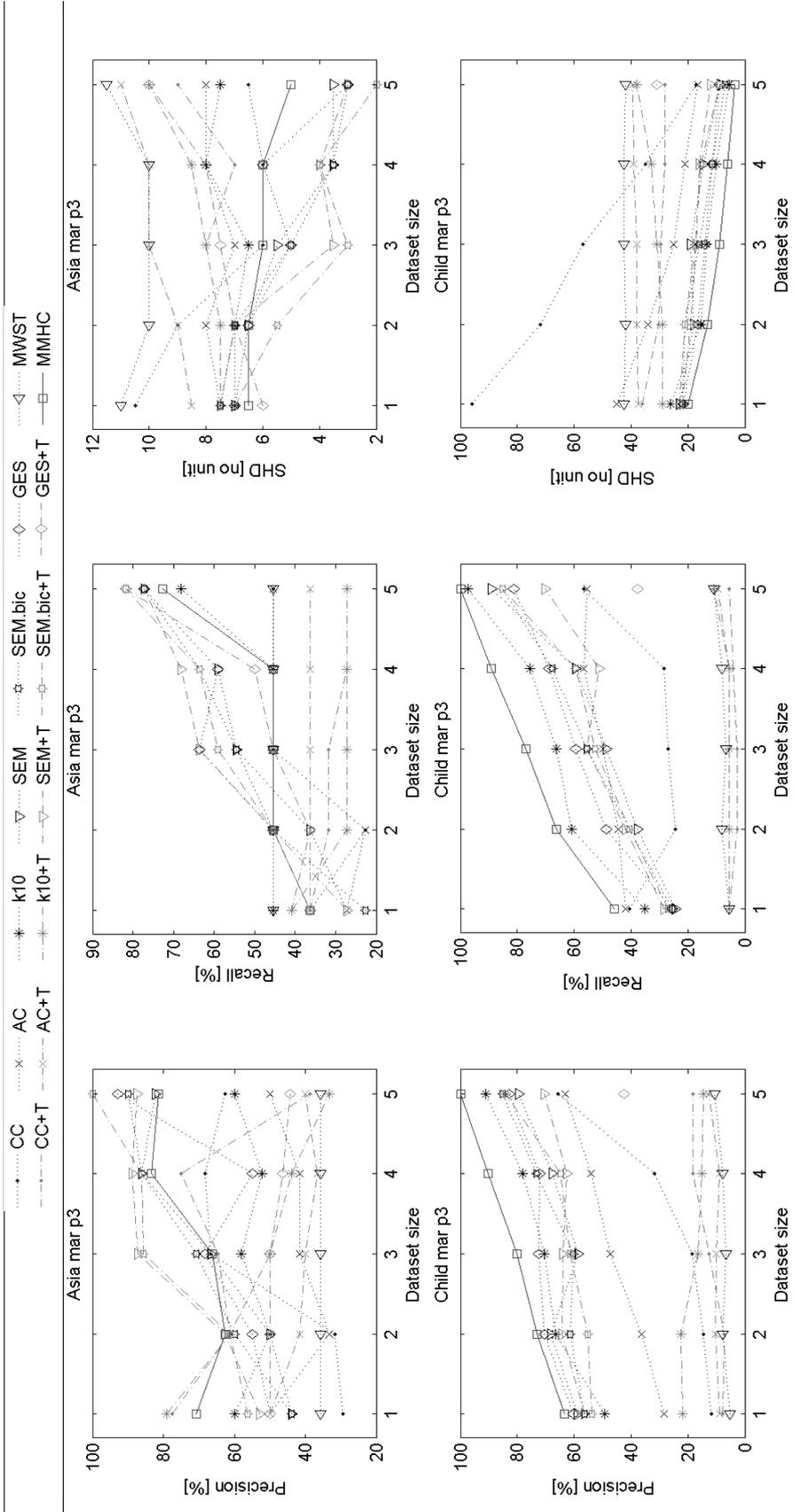


Precision-Recall-SHD Insurance e Alarm1 - mar p2

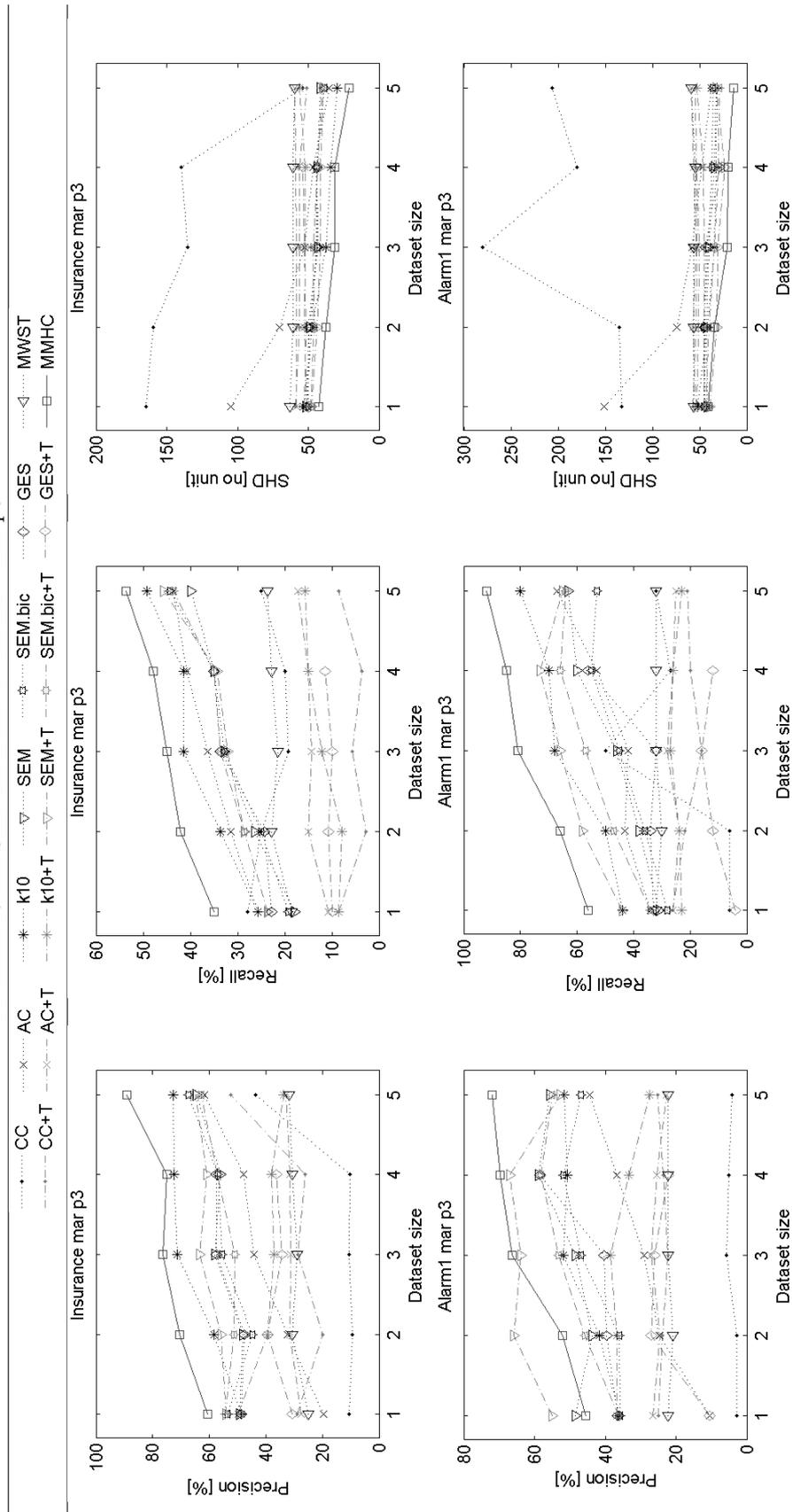




Precision-Recall-SHD Asia e Child - mar p3

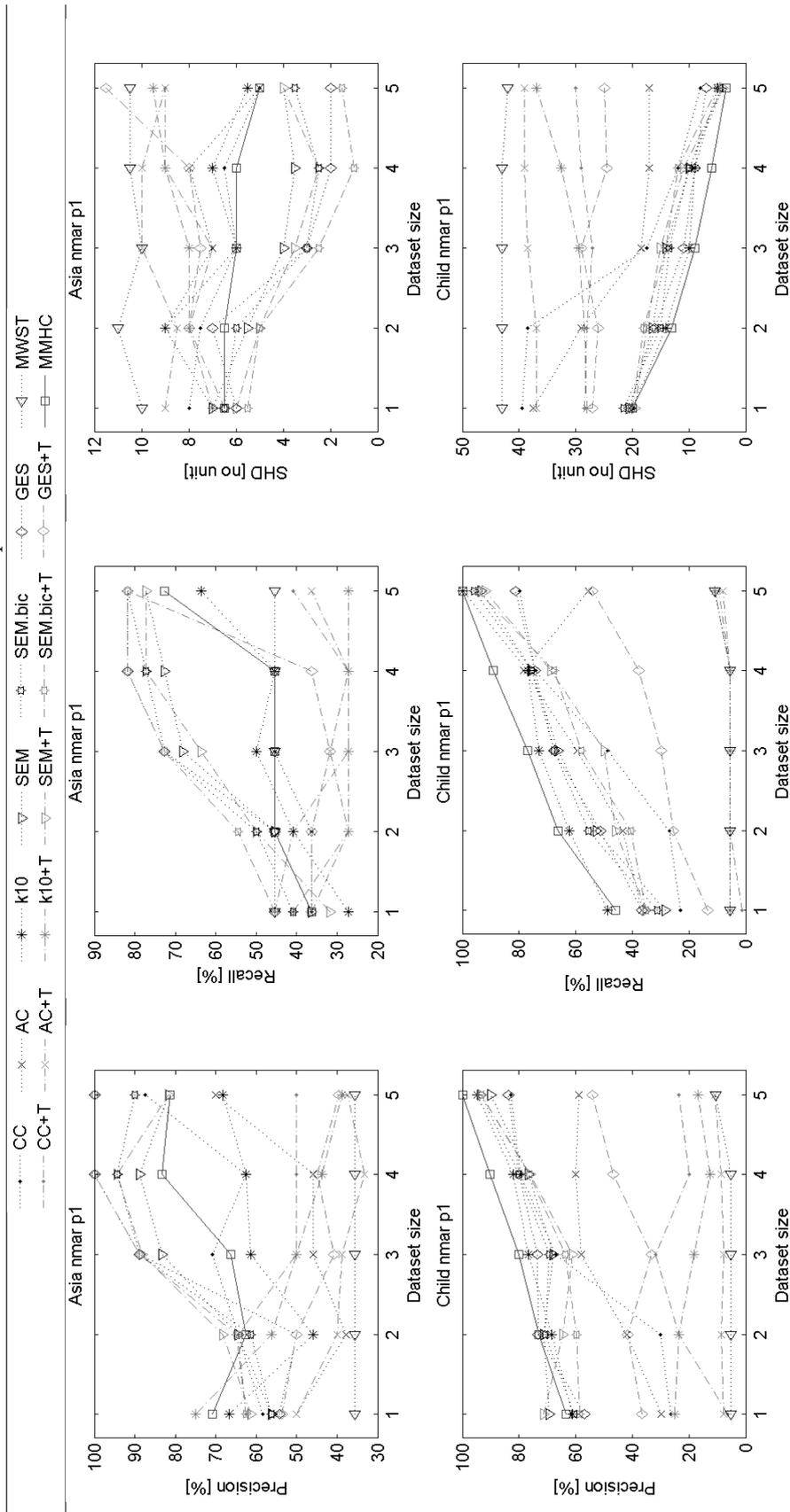


Precision-Recall-SHD Insurance e Alarm1 - mar p3

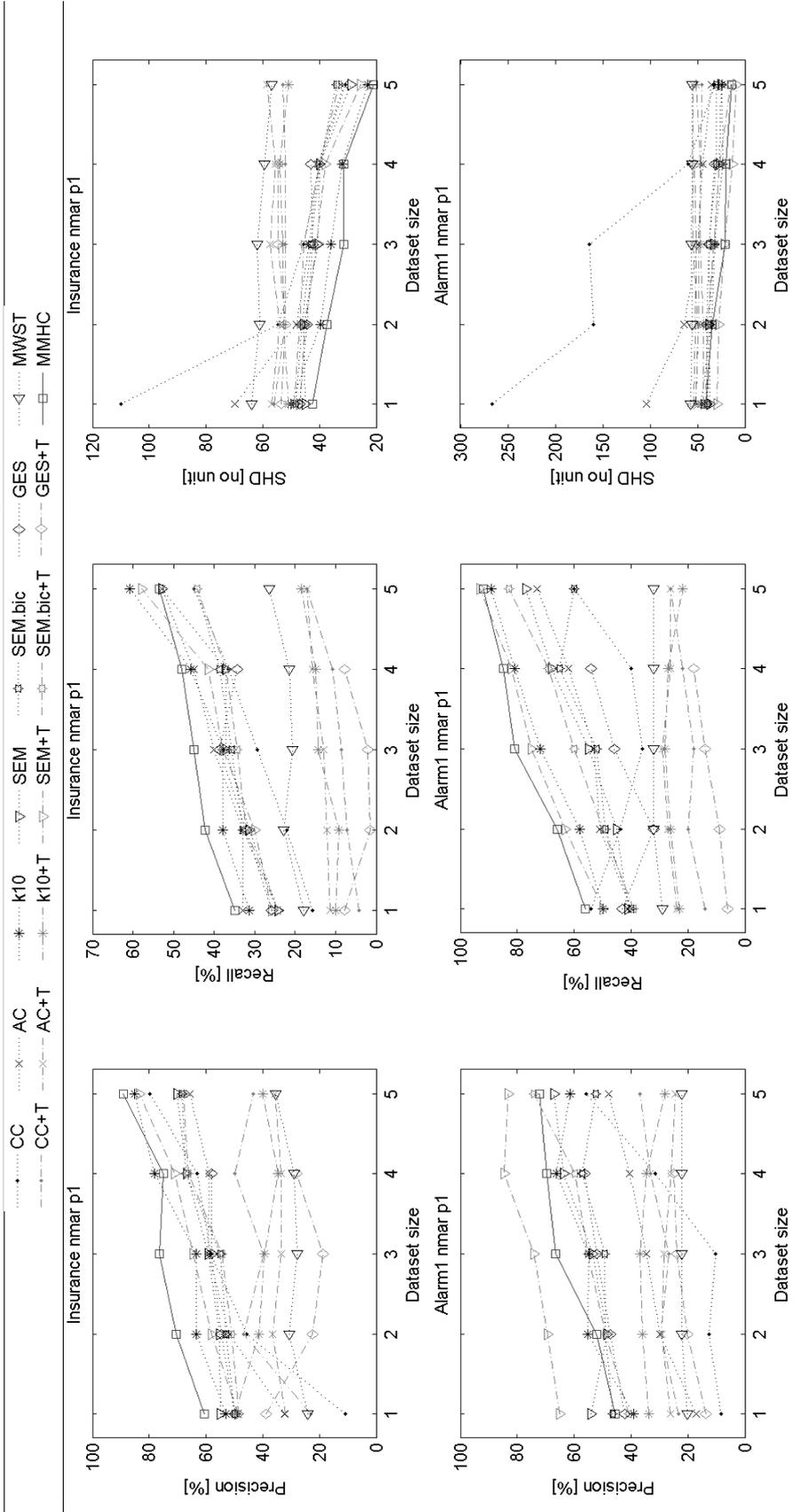




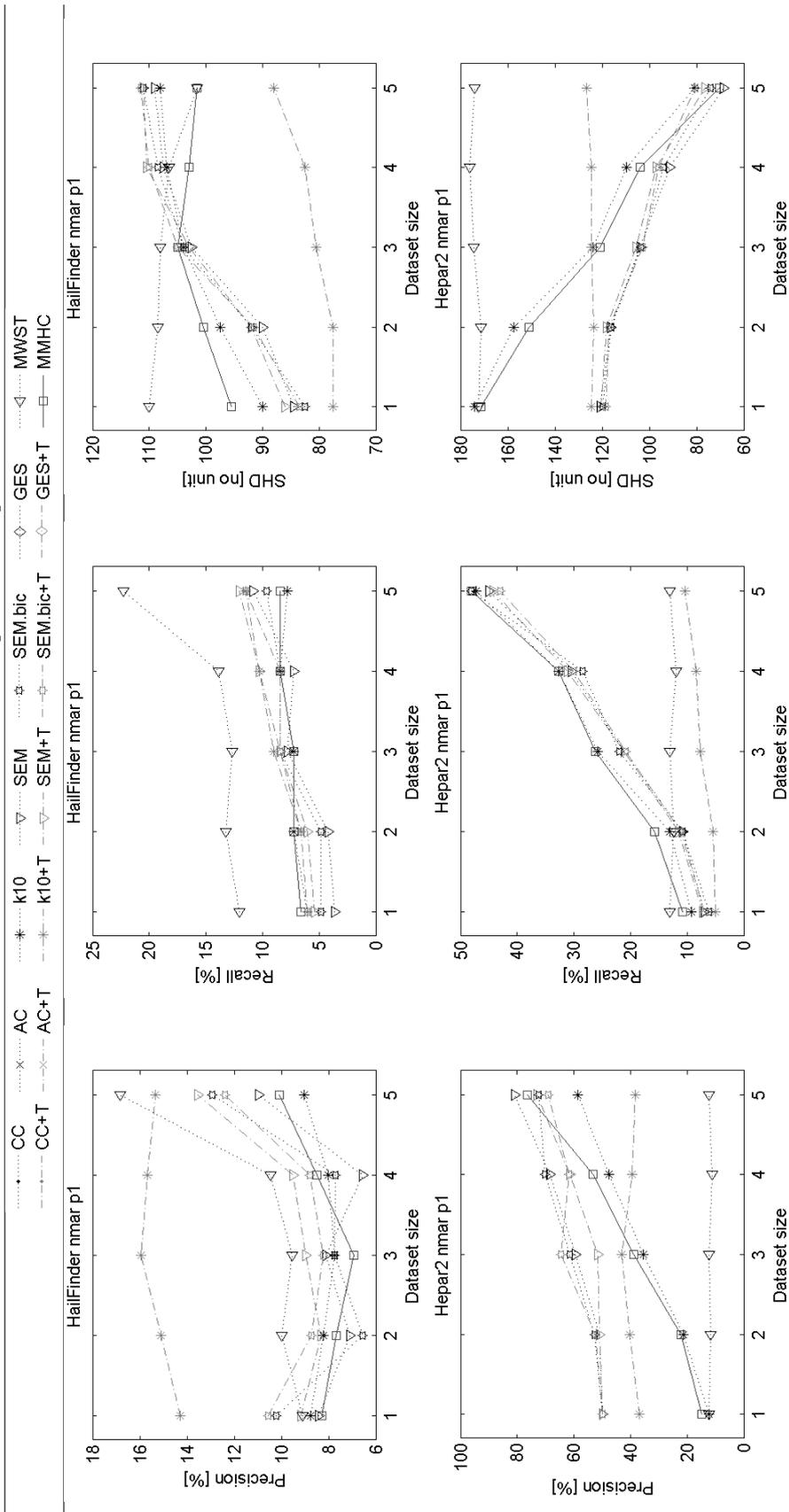
Precision-Recall-SHD Asia e Child - nmar p1



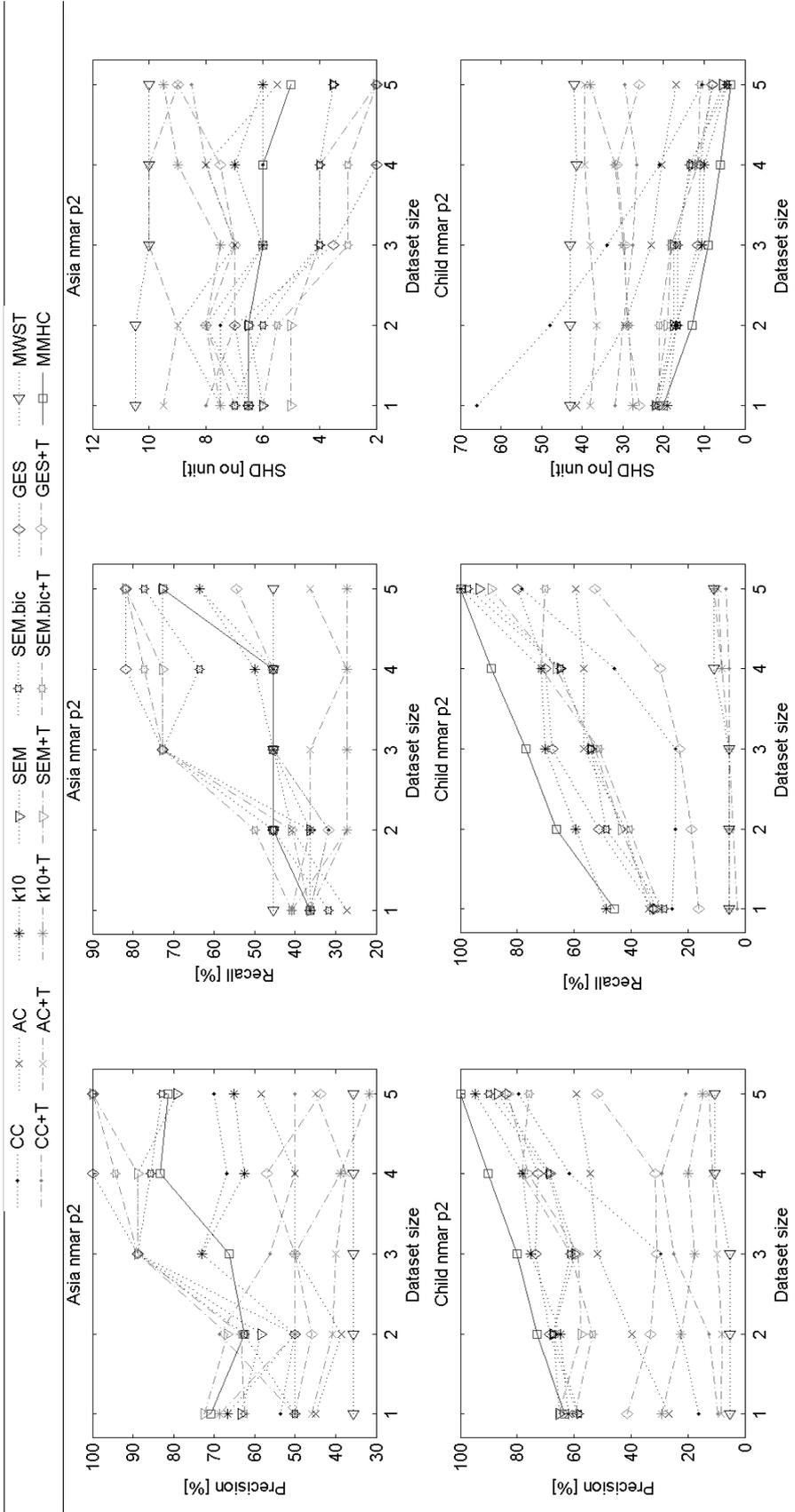
Precision-Recall-SHD Insurance e Alarm1 - nmar p1



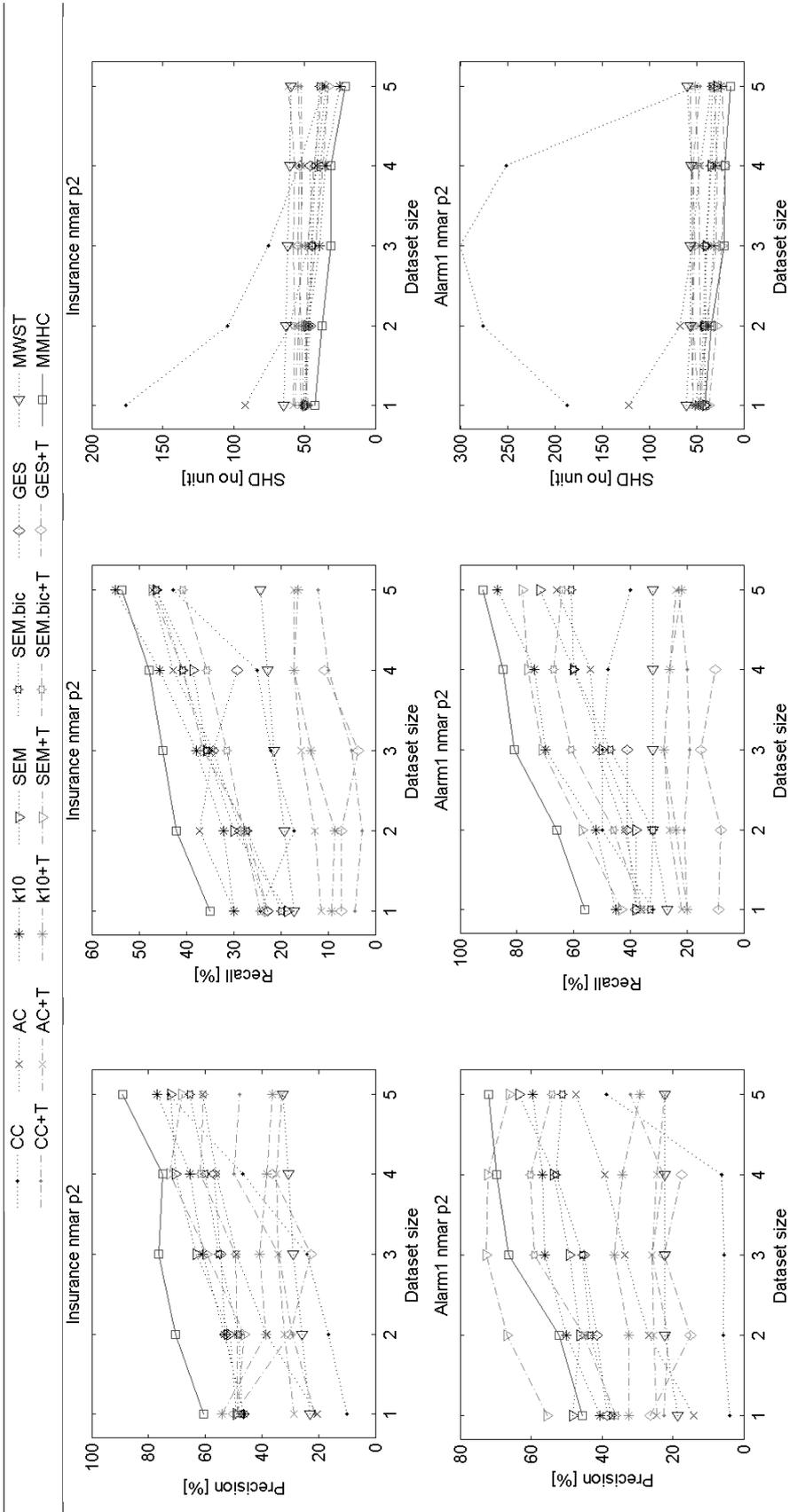
Precision-Recall-SHD HailFinder e Hepar2 - nmar p1



Precision-Recall-SHD Asia e Child - nmar p2

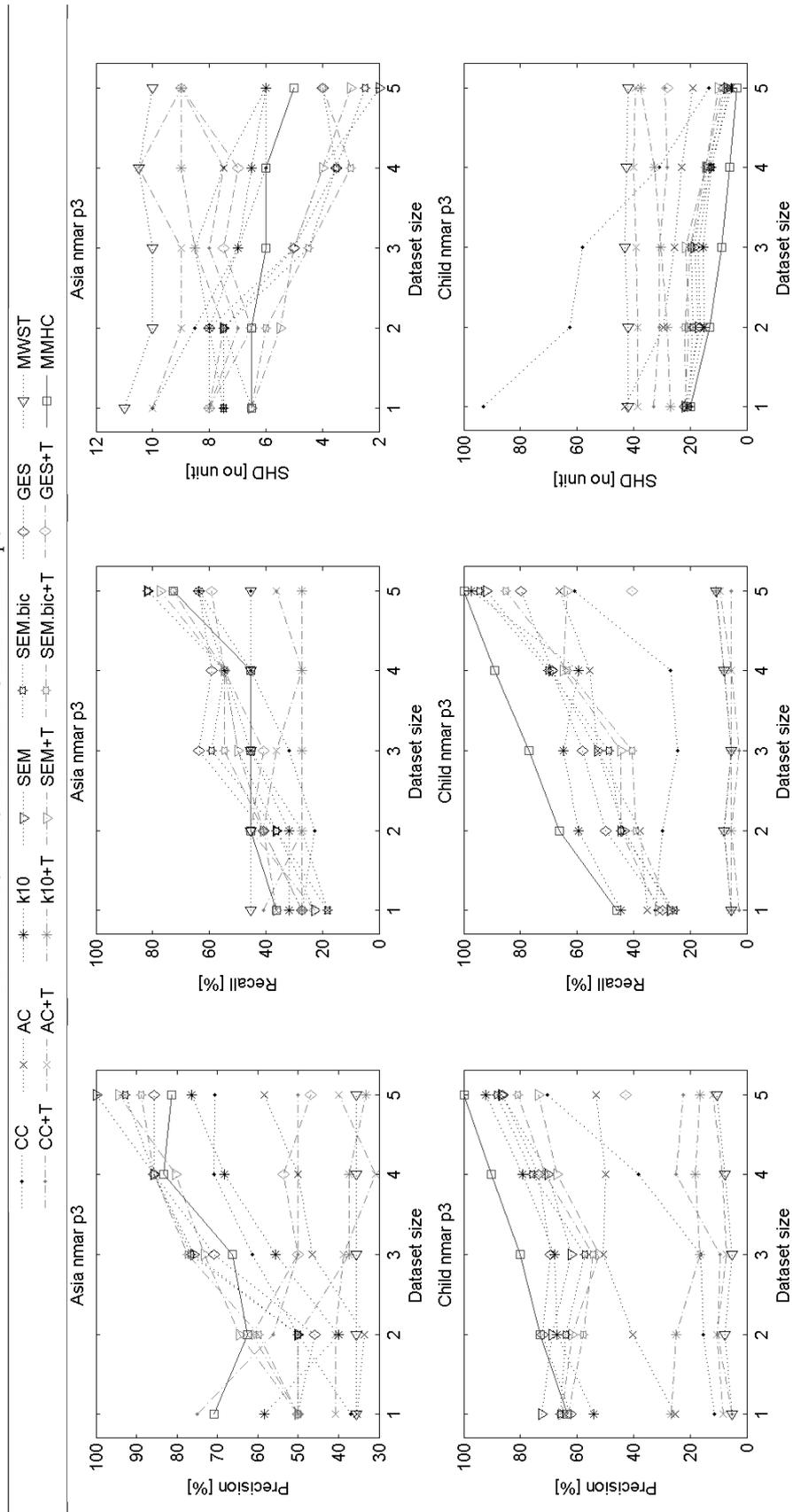


Precision-Recall-SHD Insurance e Alarm1 - nmar p2

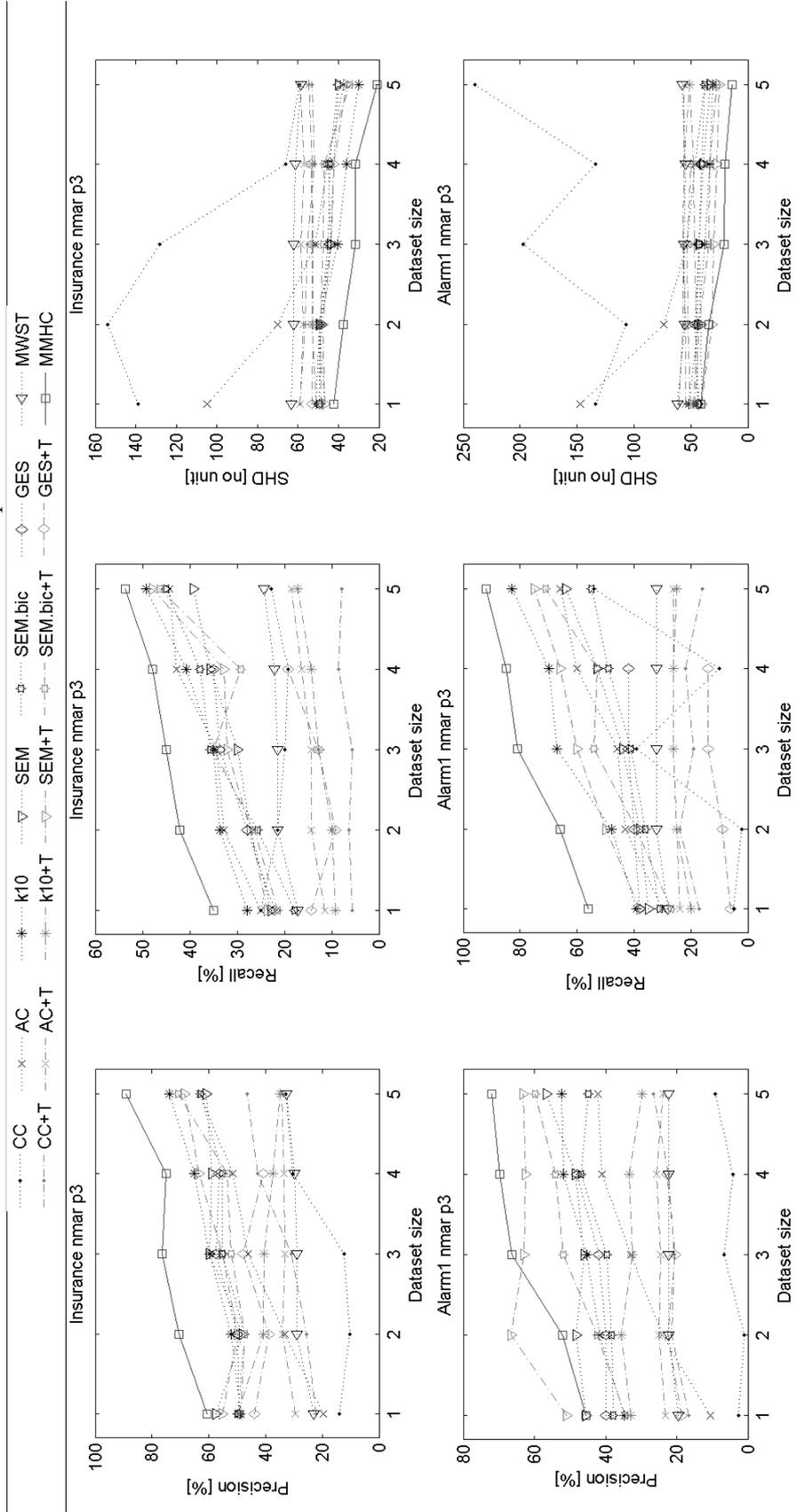




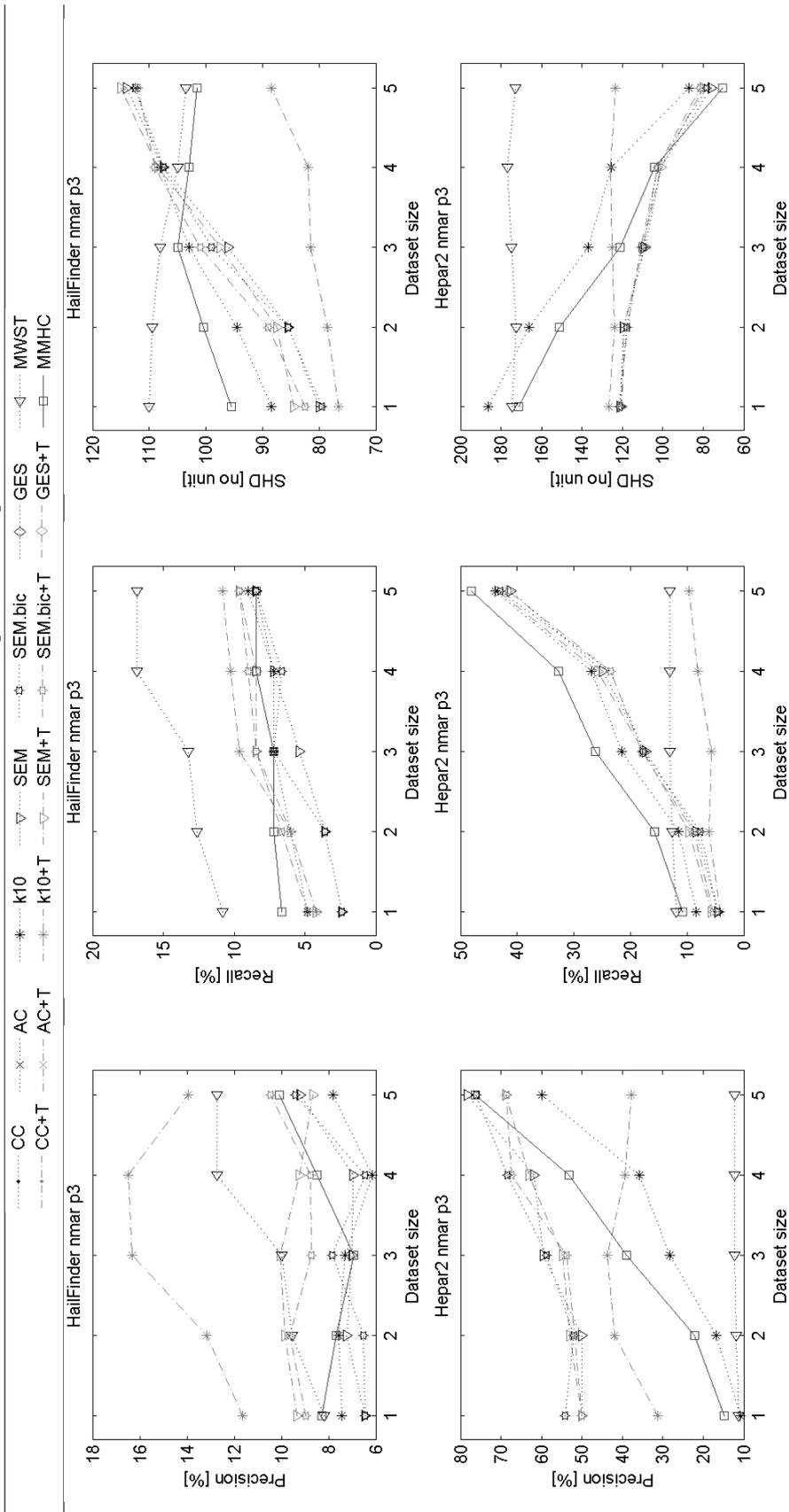
Precision-Recall-SHD Asia e Child - nmar p3



Precision-Recall-SHD Insurance e Alarm1 - nmar p3

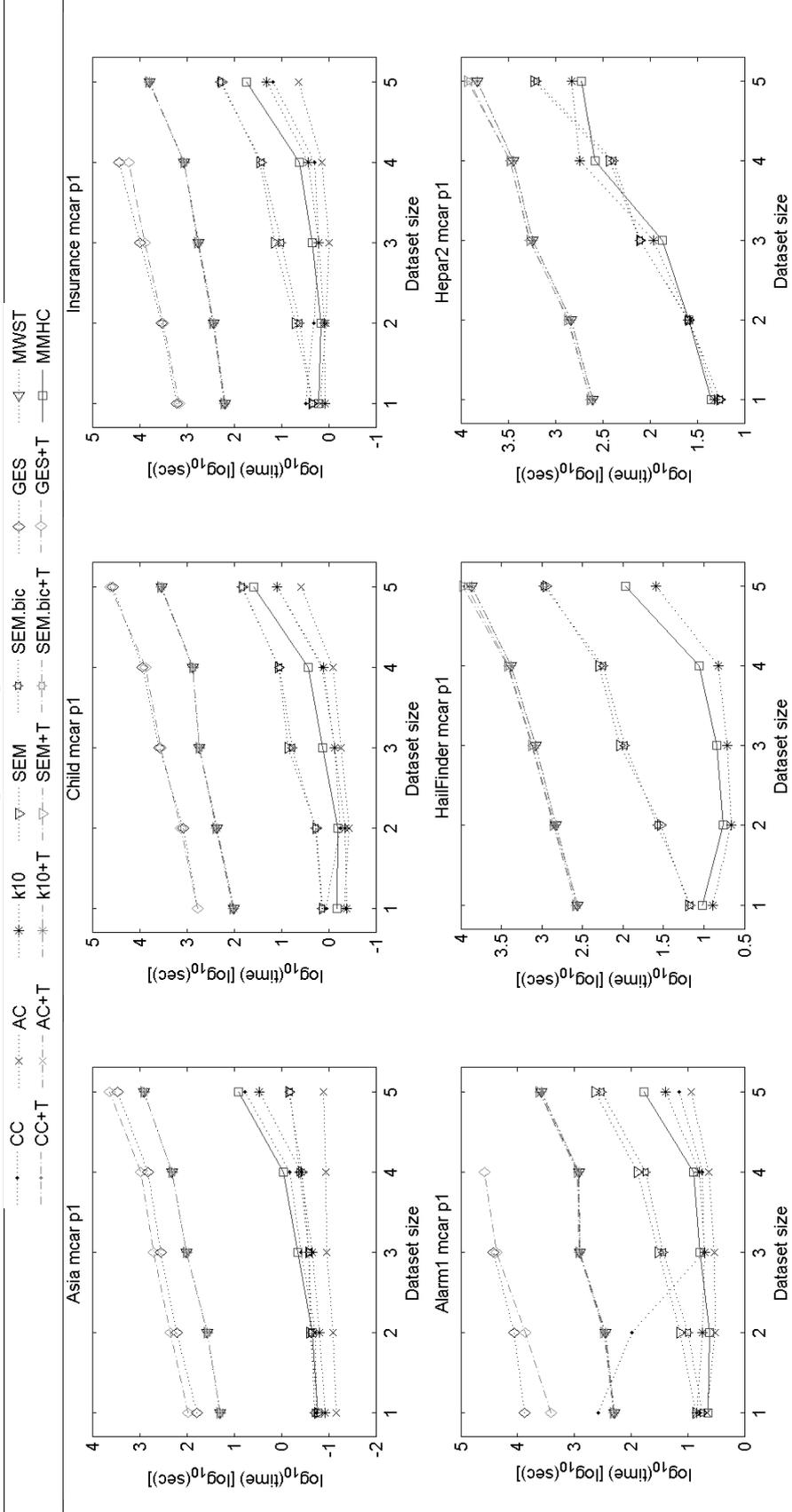


Precision-Recall-SHD HailFinder e Hepar2 - nmar p3

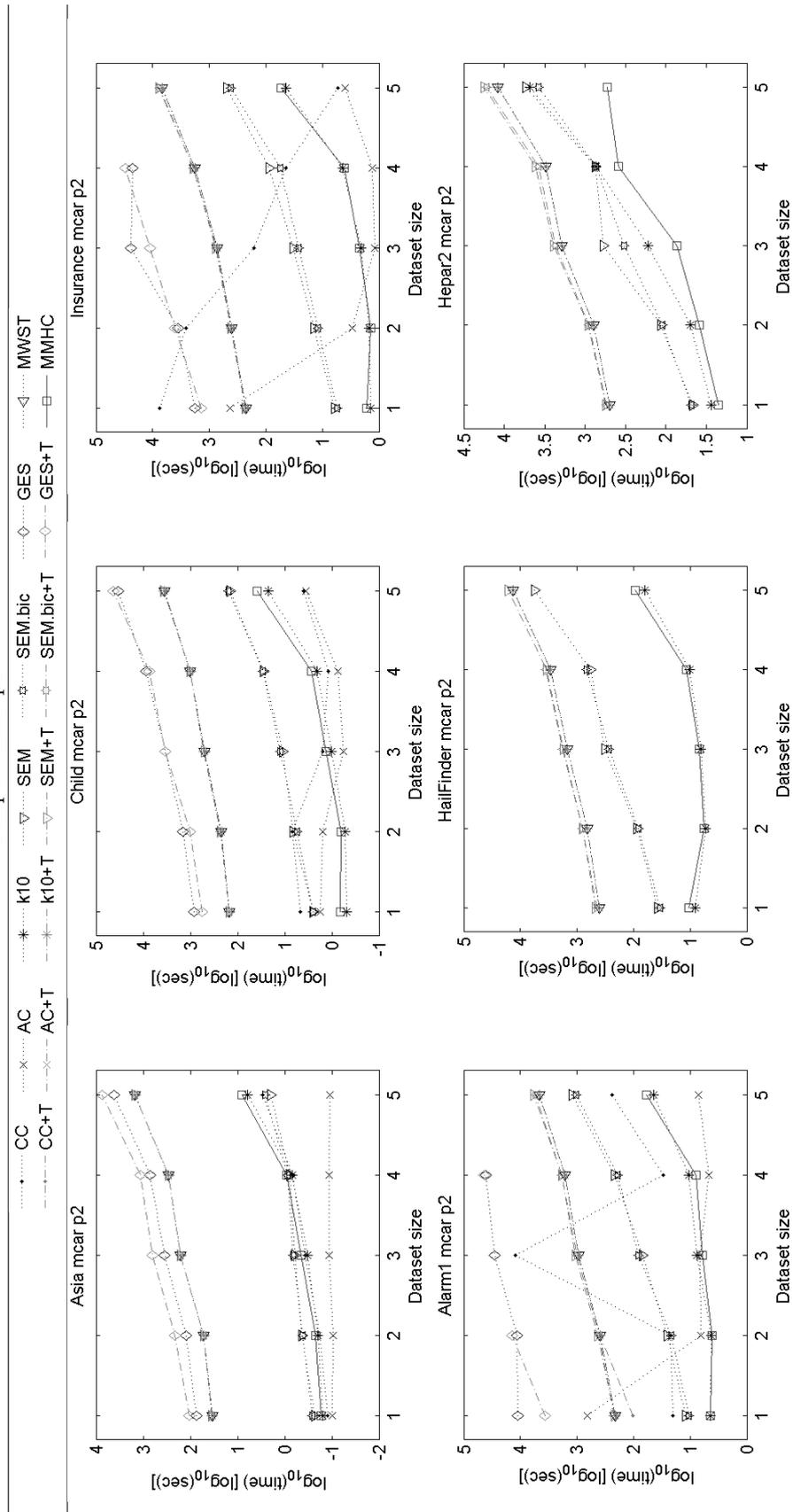


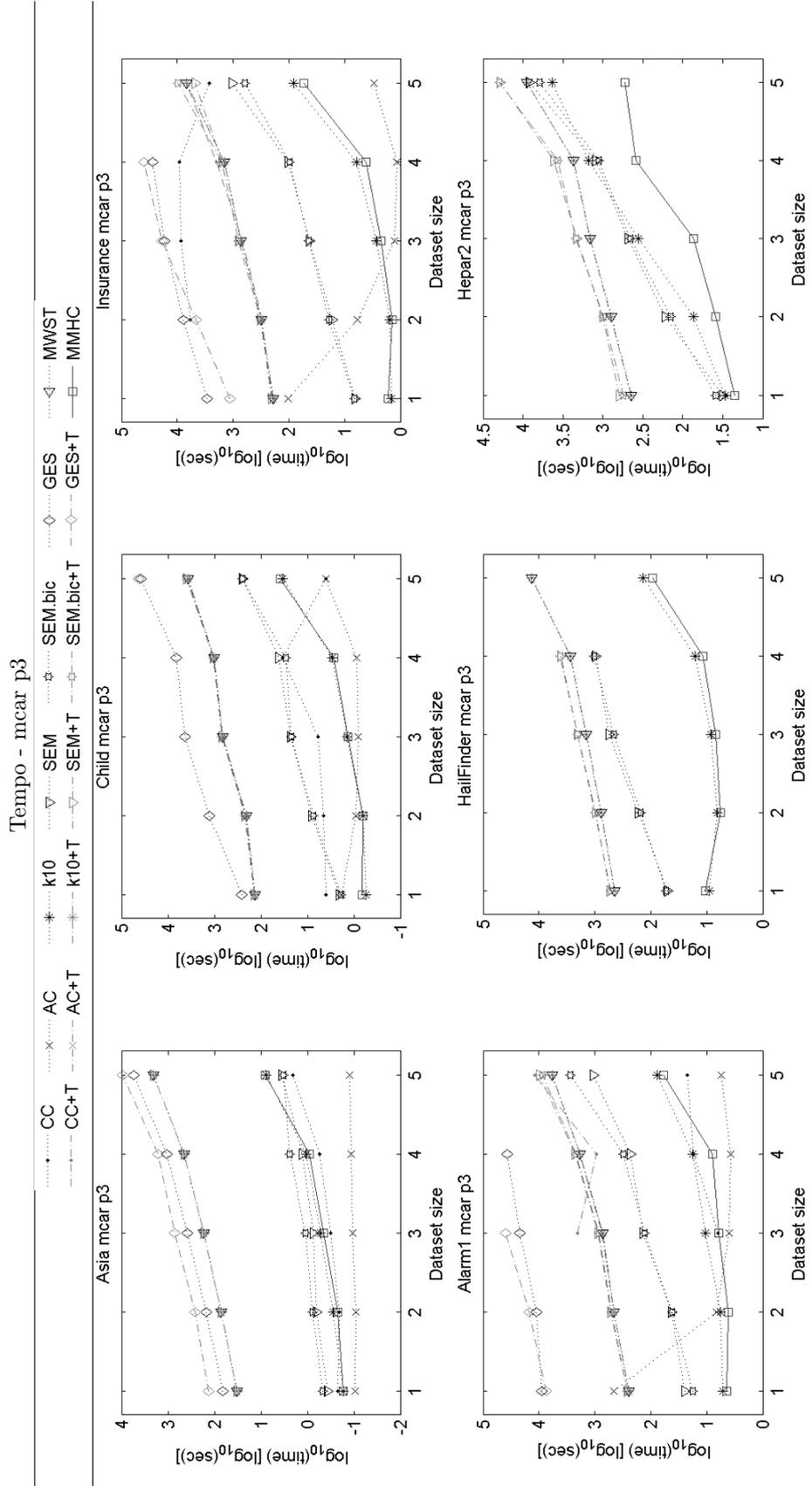
Tempo

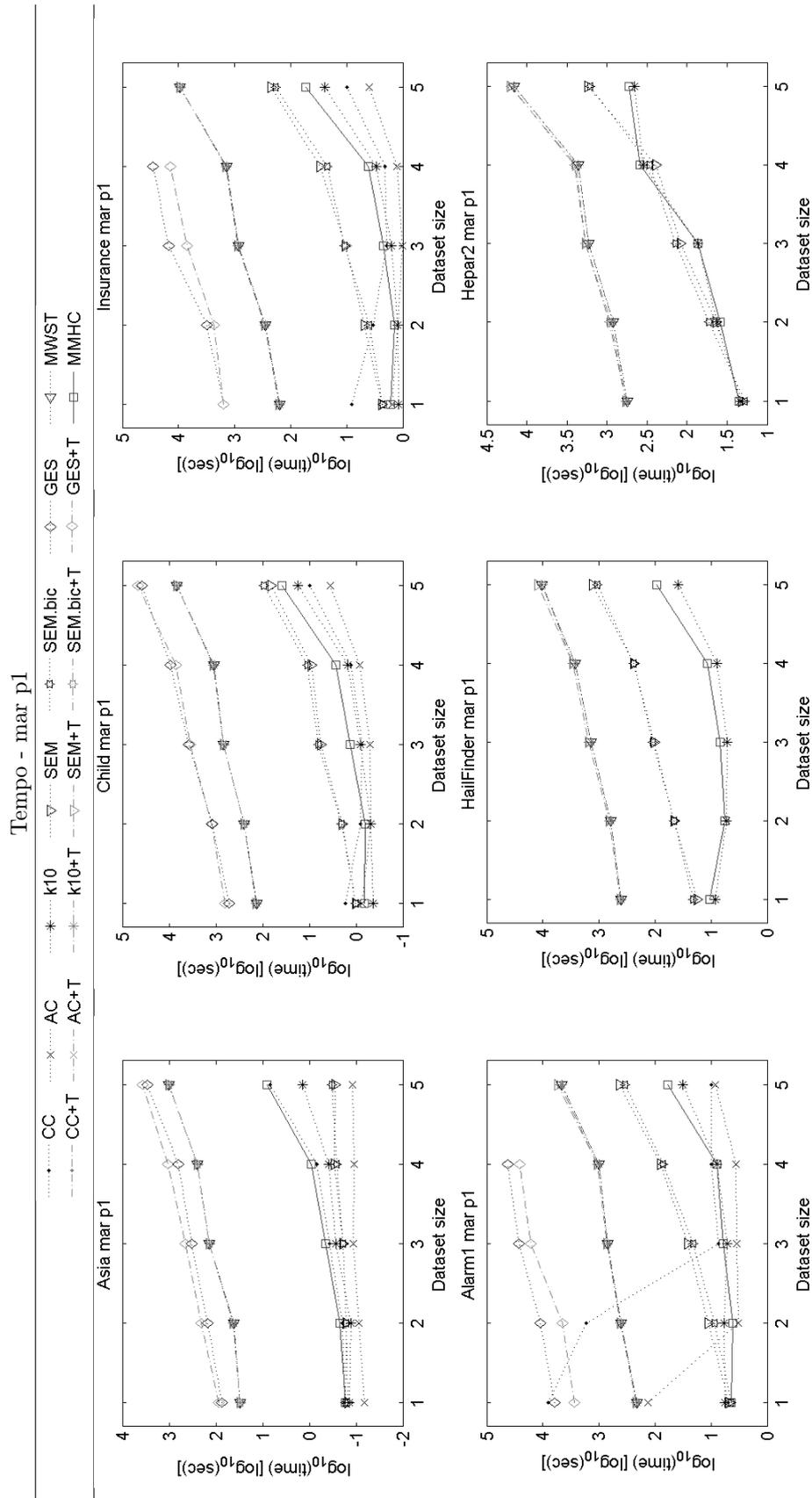
Tempo - mcar p1

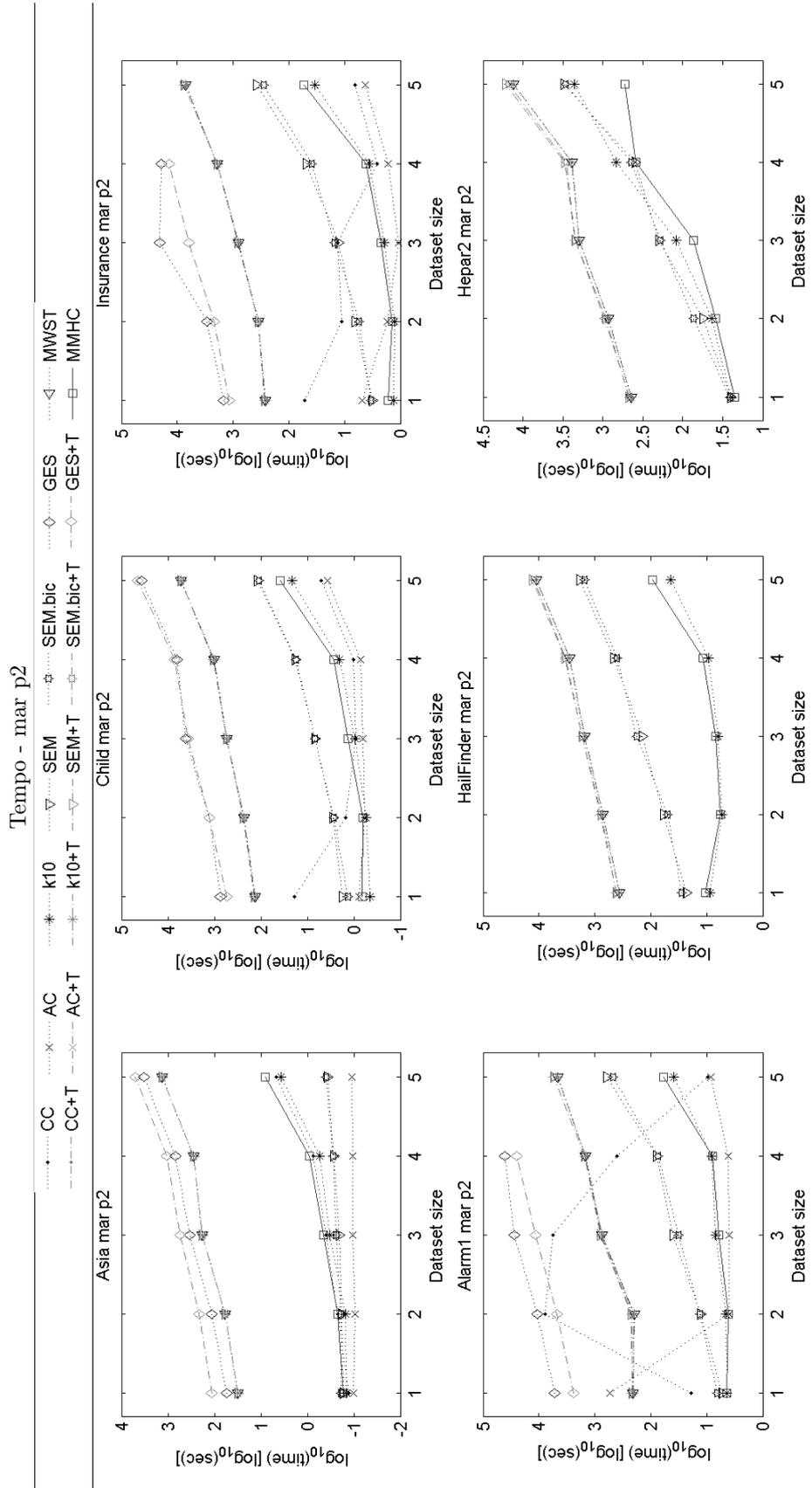


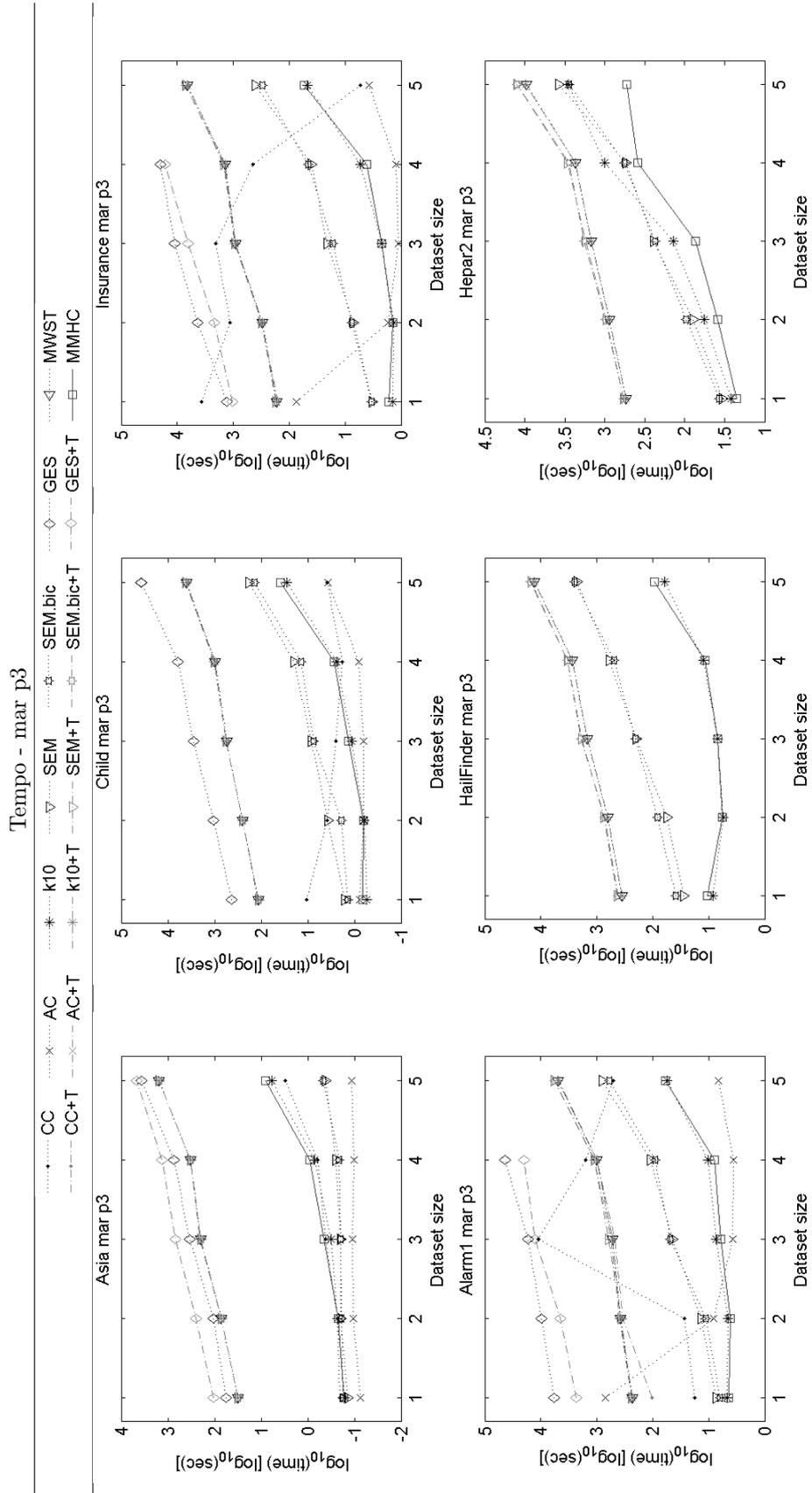
Tempo - mcar p2



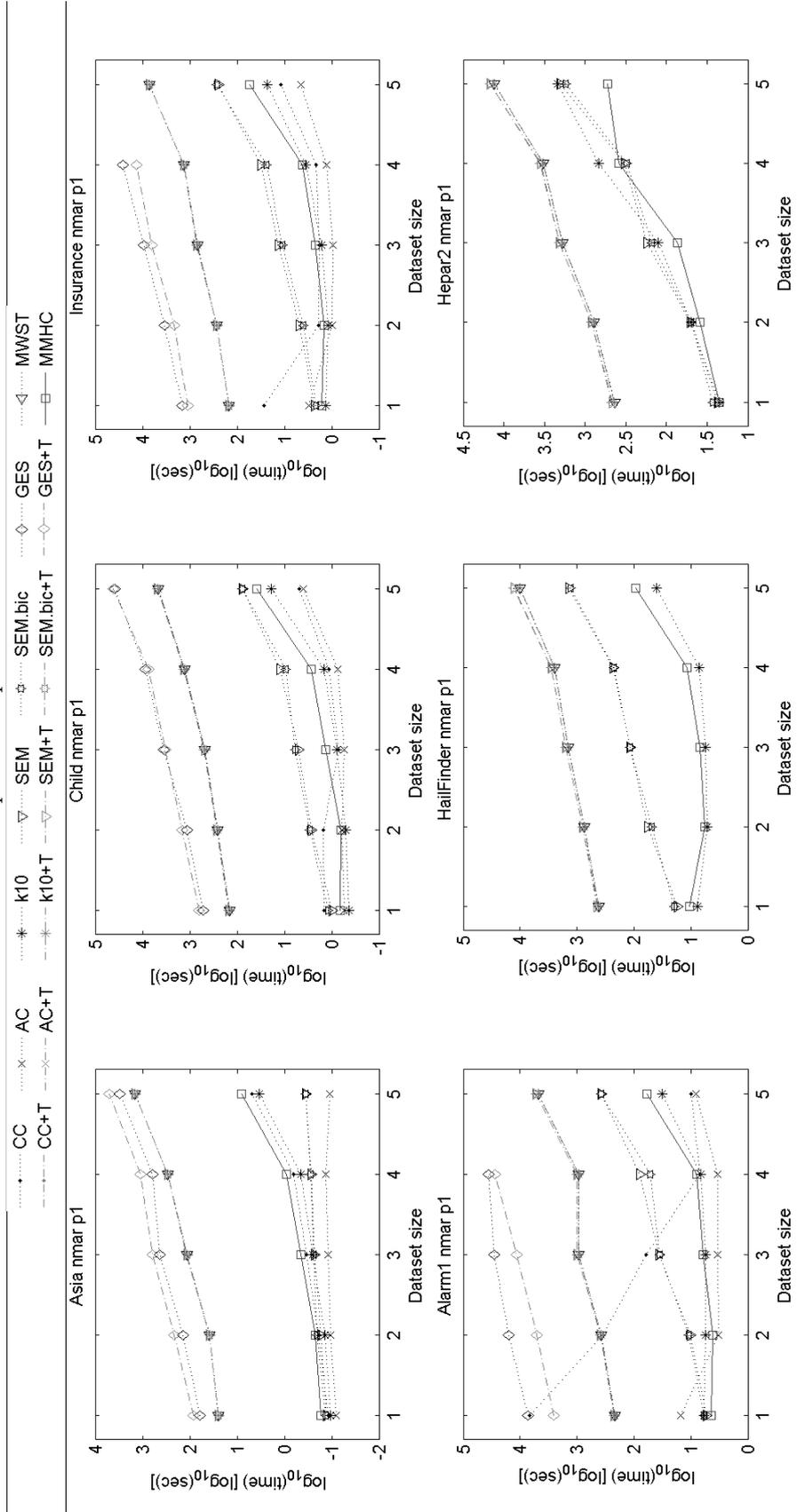


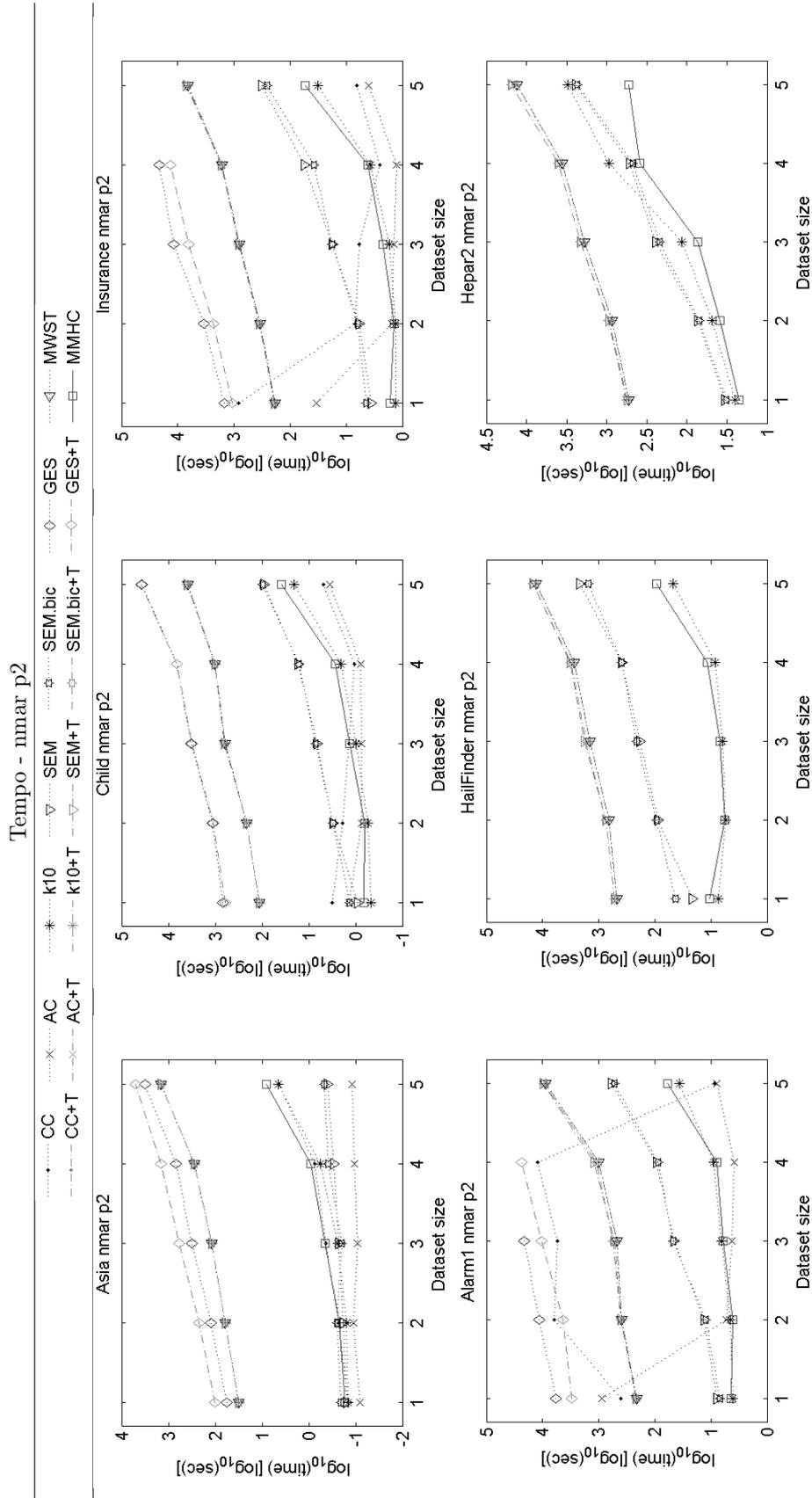




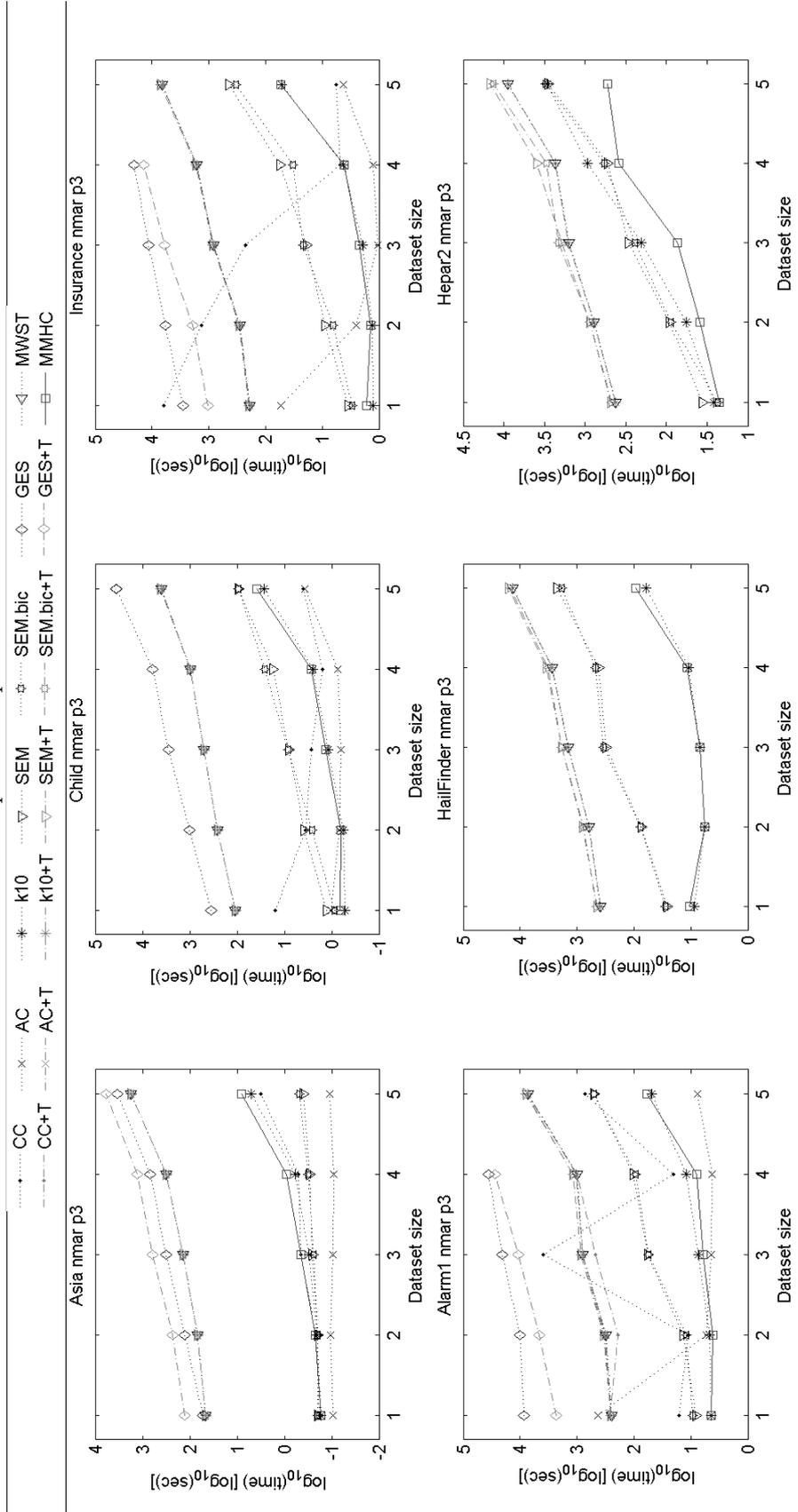


Tempo - nmar p1



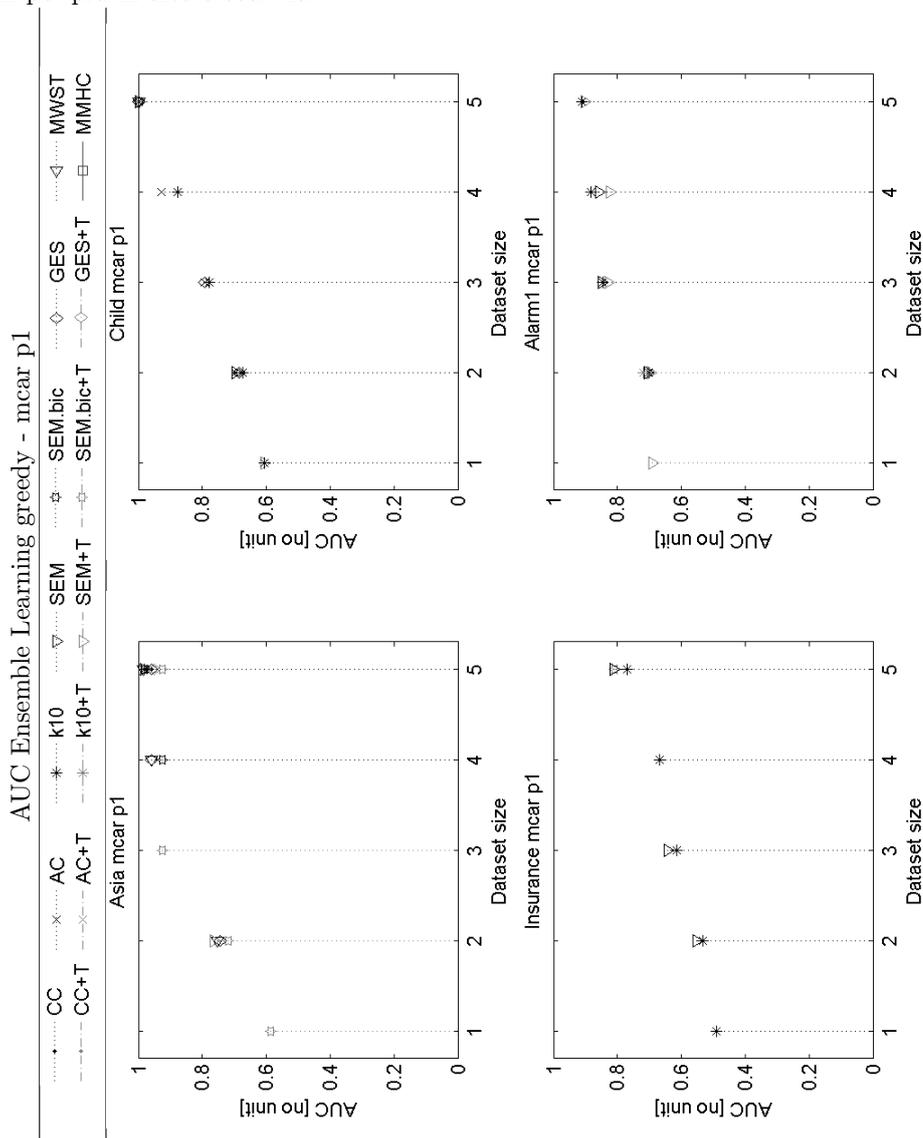


Tempo - nmar p3

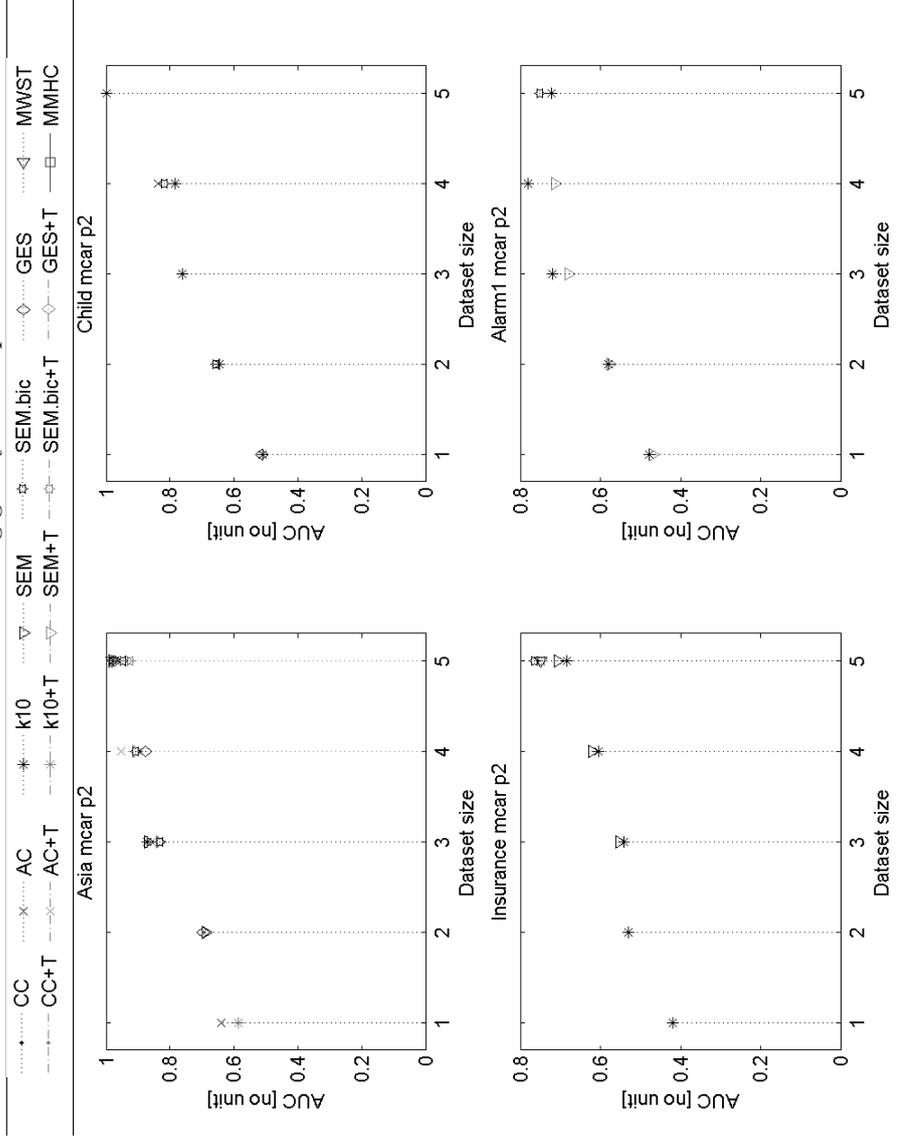


**AUC greedy**

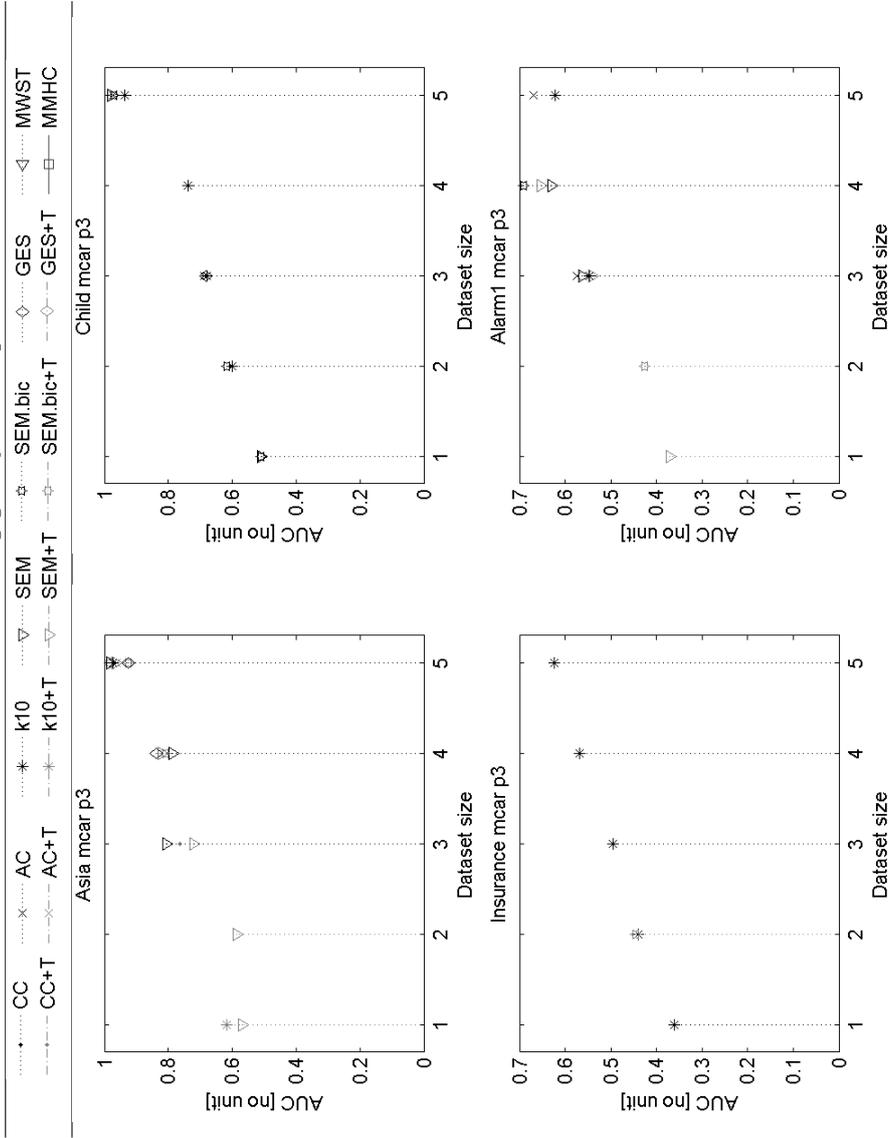
Poichè si è usato un approccio greedy, i grafici che seguono indicano i metodi selezionati per ogni caso mostrando l'aggiunta di un algoritmo solo se questo migliora le performance in termini di AUC della curva PvsR. Pertanto, il metodo che da solo ha una AUC maggiore è quello più in basso, il secondo che la migliora è un po' più in alto e così via.



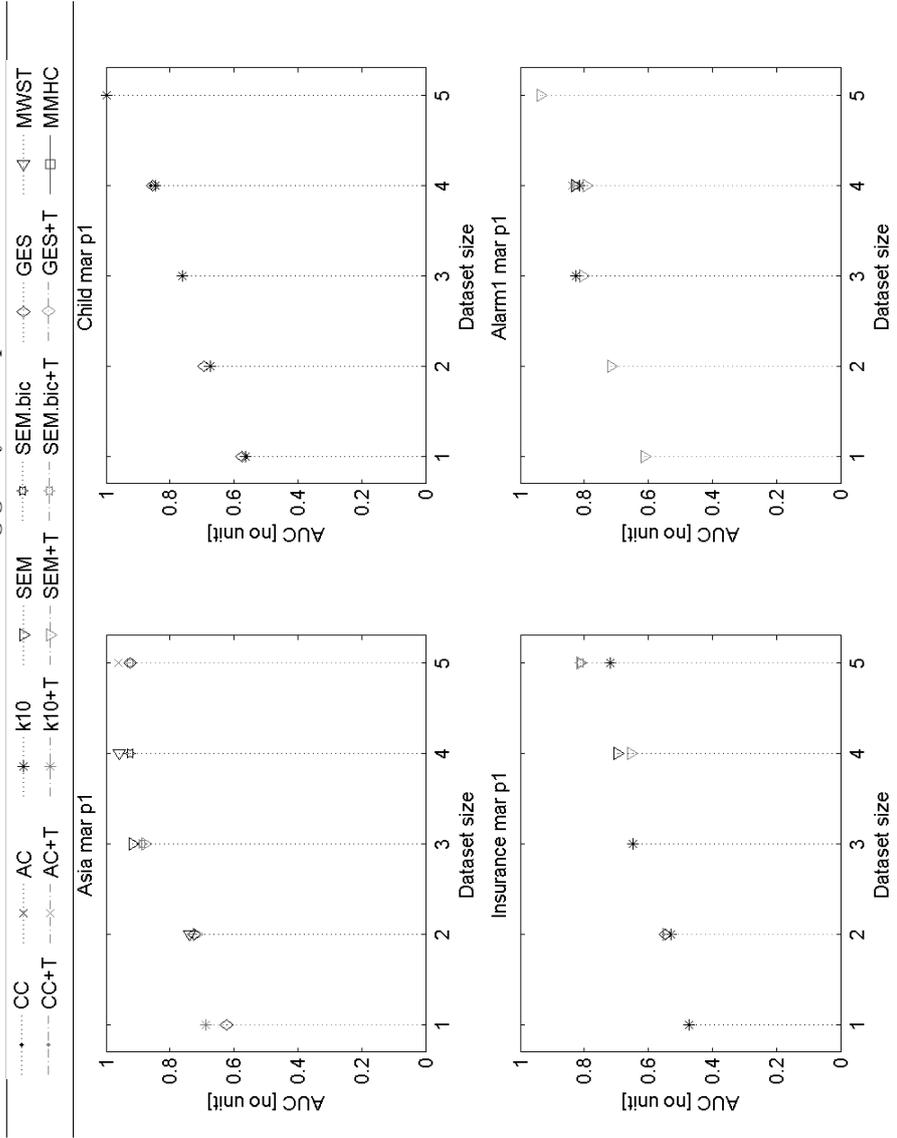
AUC Ensemble Learning greedy - mcar p2



AUC Ensemble Learning greedy - mcar p3

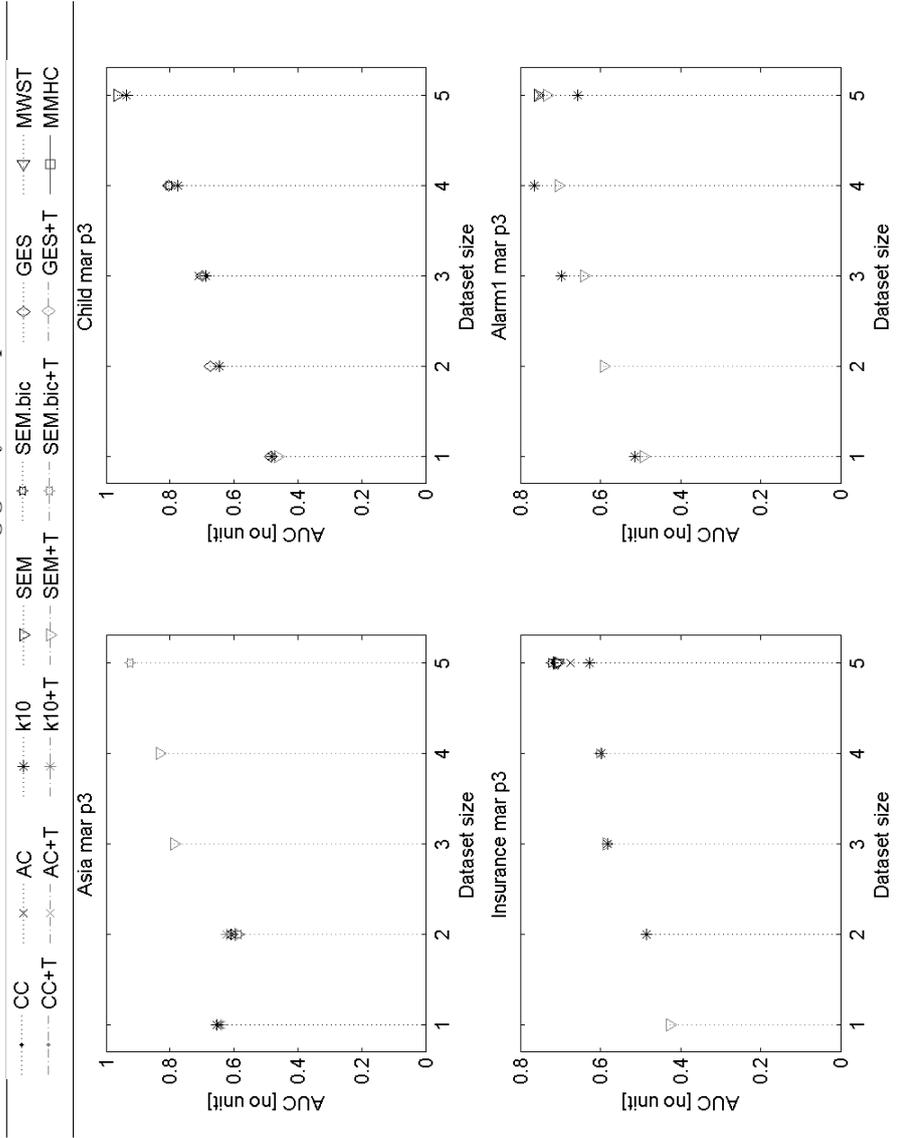


AUC Ensemble Learning greedy - mar p1

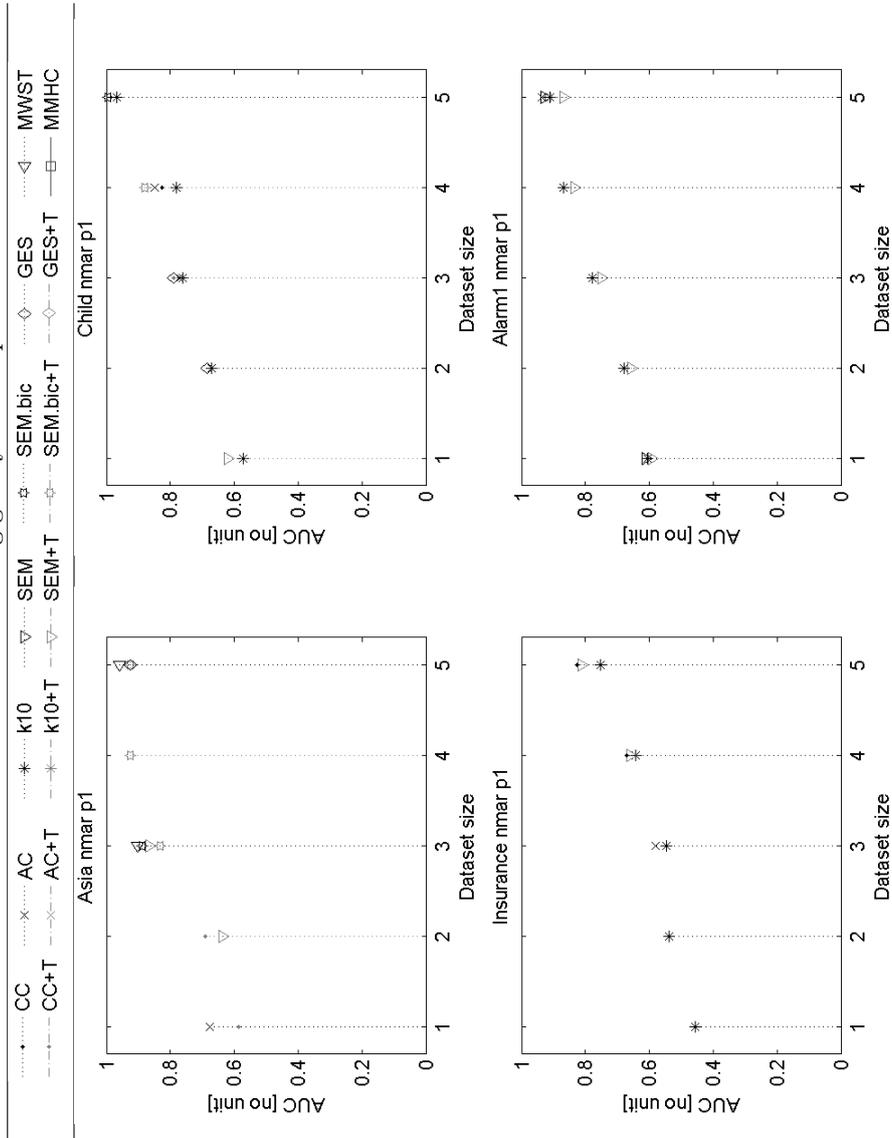




AUC Ensemble Learning greedy - mar p3



AUC Ensemble Learning greedy - nmar p1

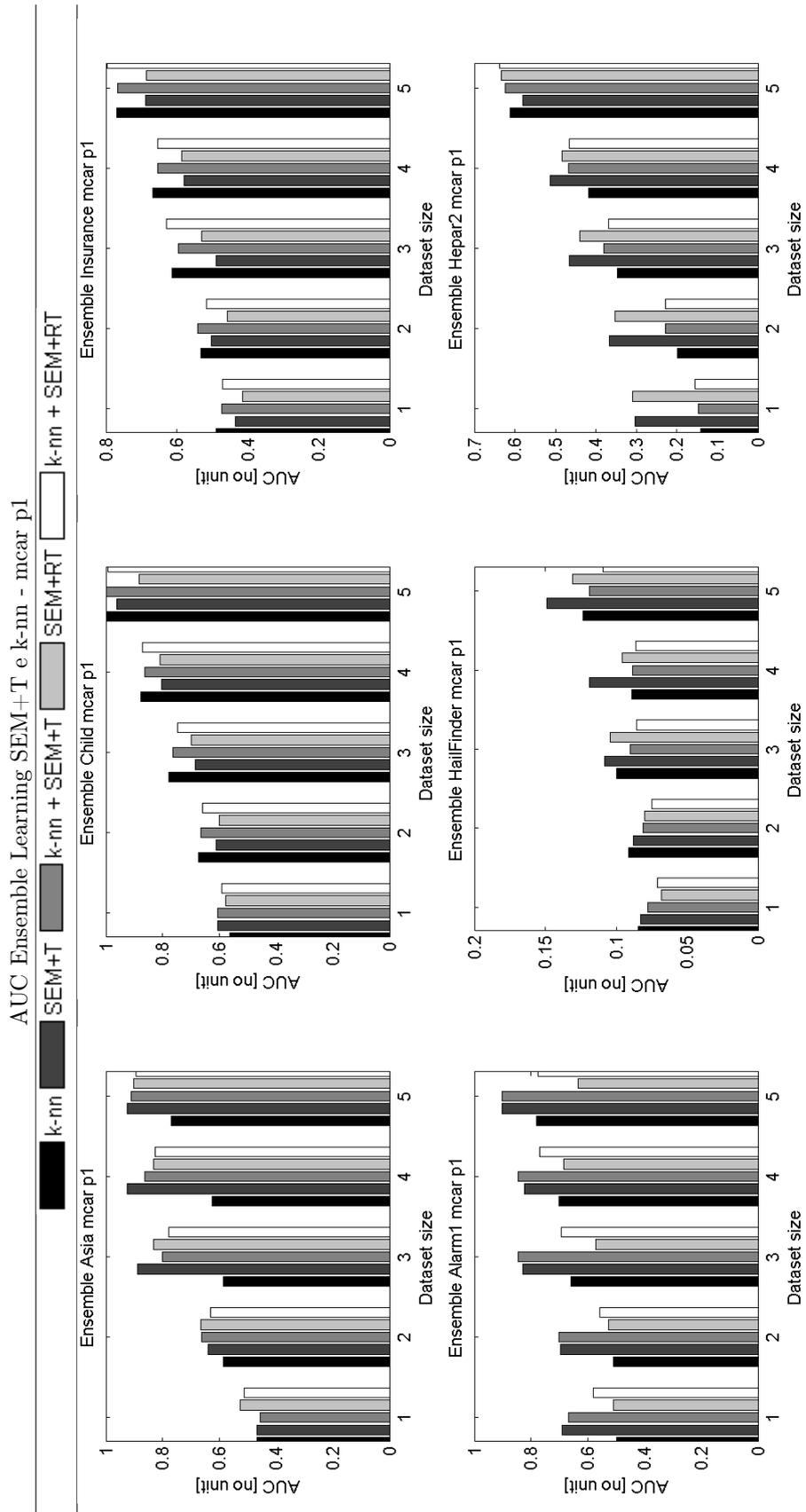




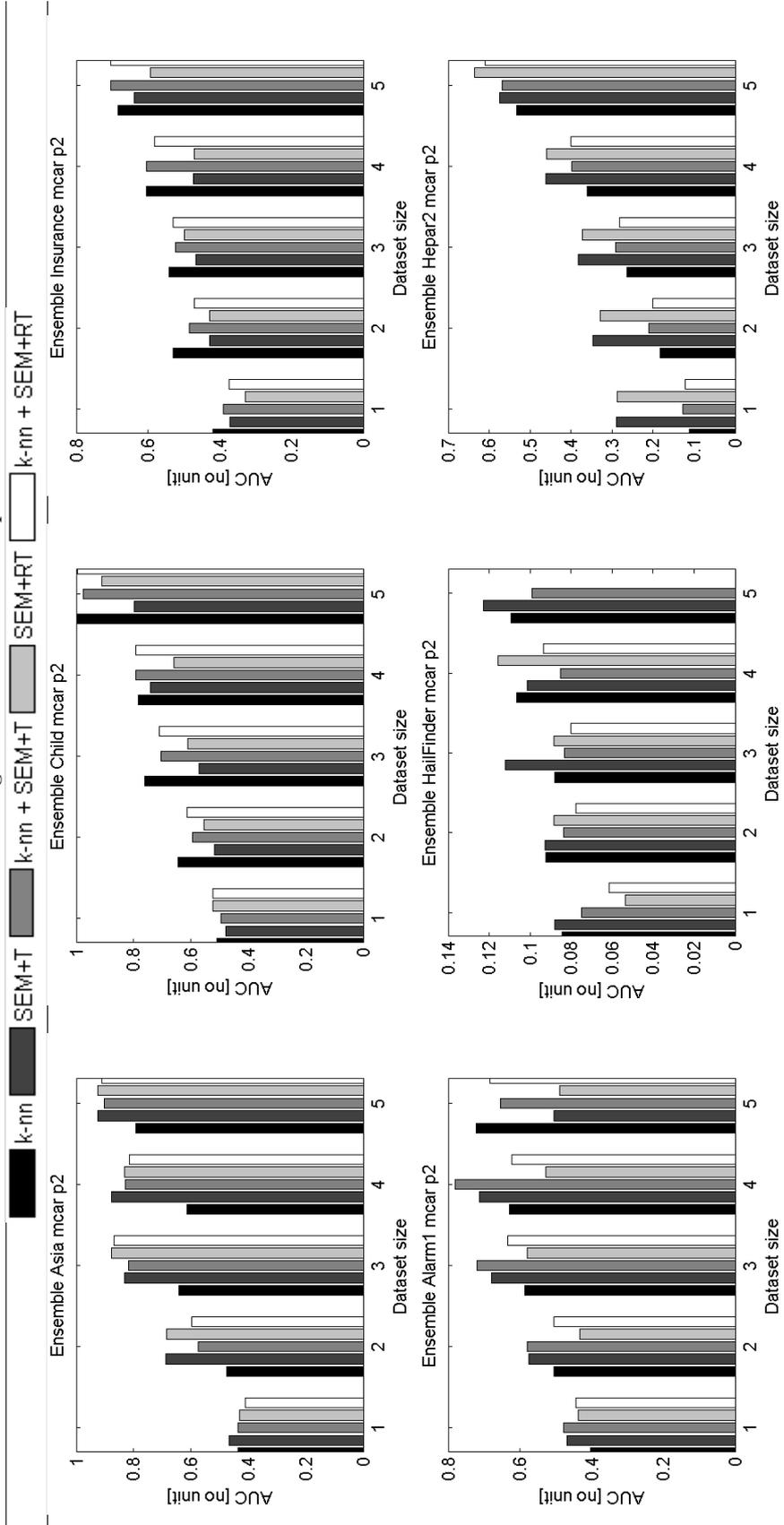


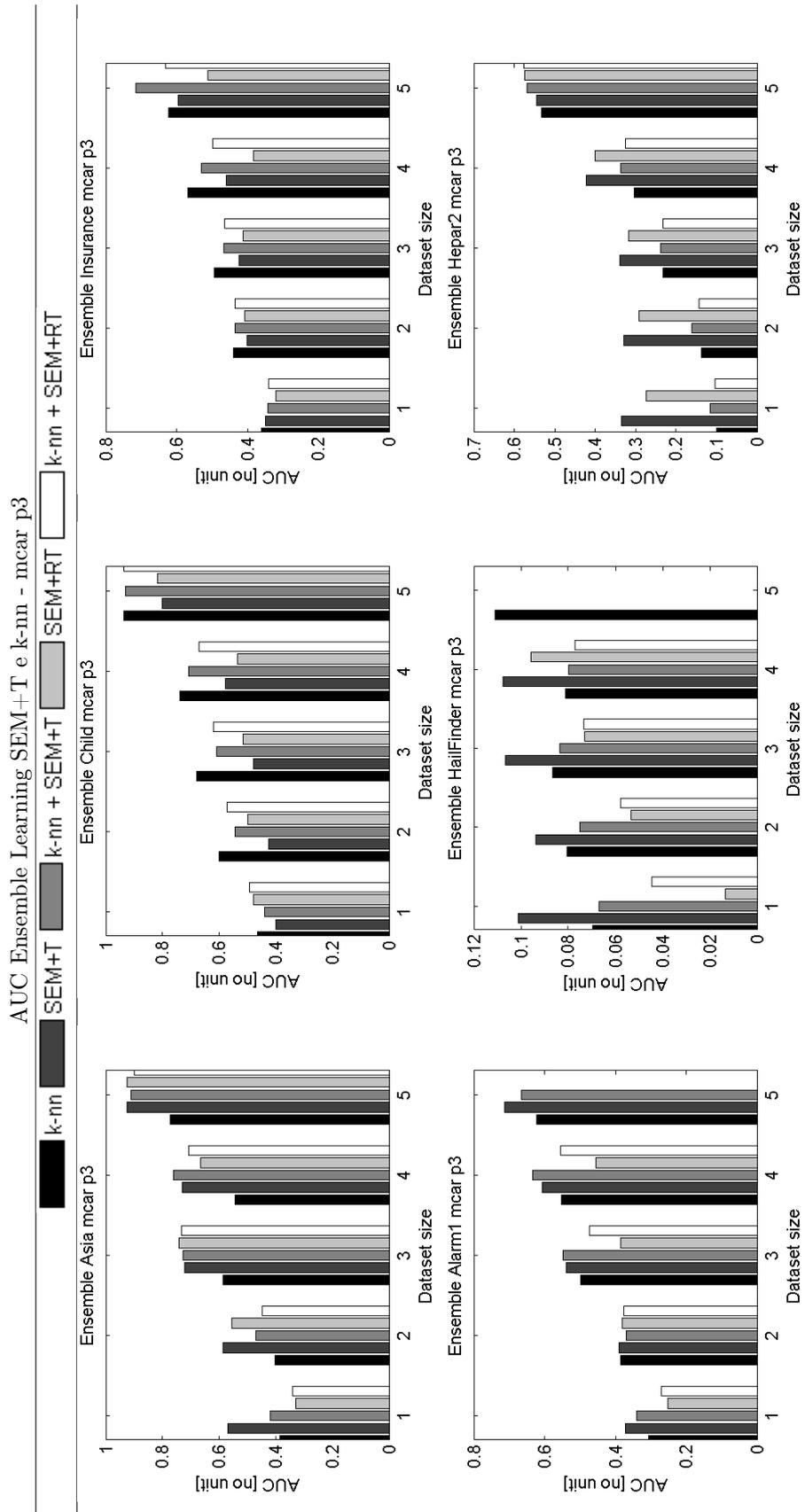
**AUC k-nn e SEM+T**

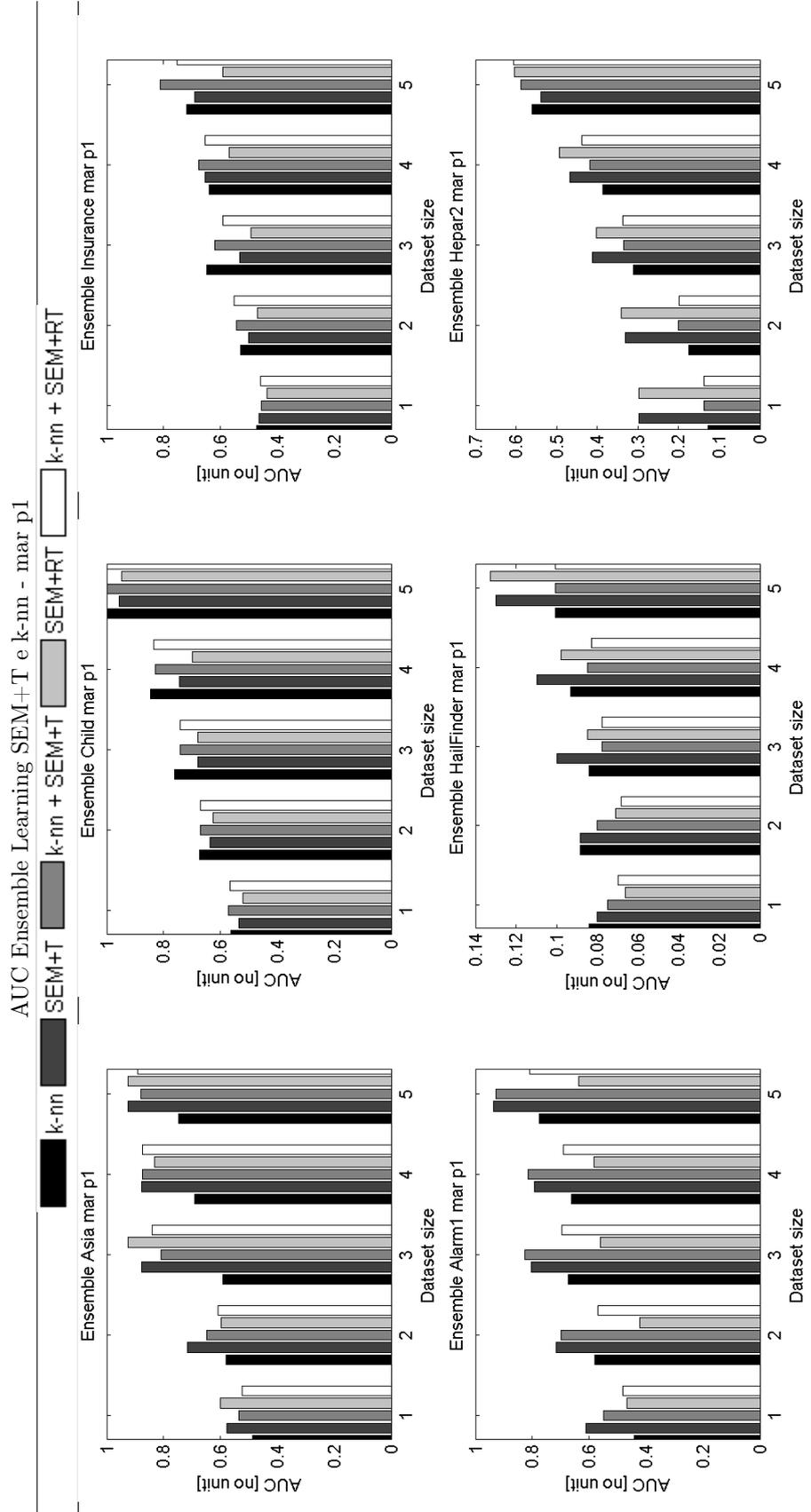
I grafici seguenti mostrano le AUC delle curve Precision vs. Recall per i singoli metodi k-nn, SEM+T e SEM+RT e per le somme di k-nn e SEM+T e di k-nn e SEM+RT.

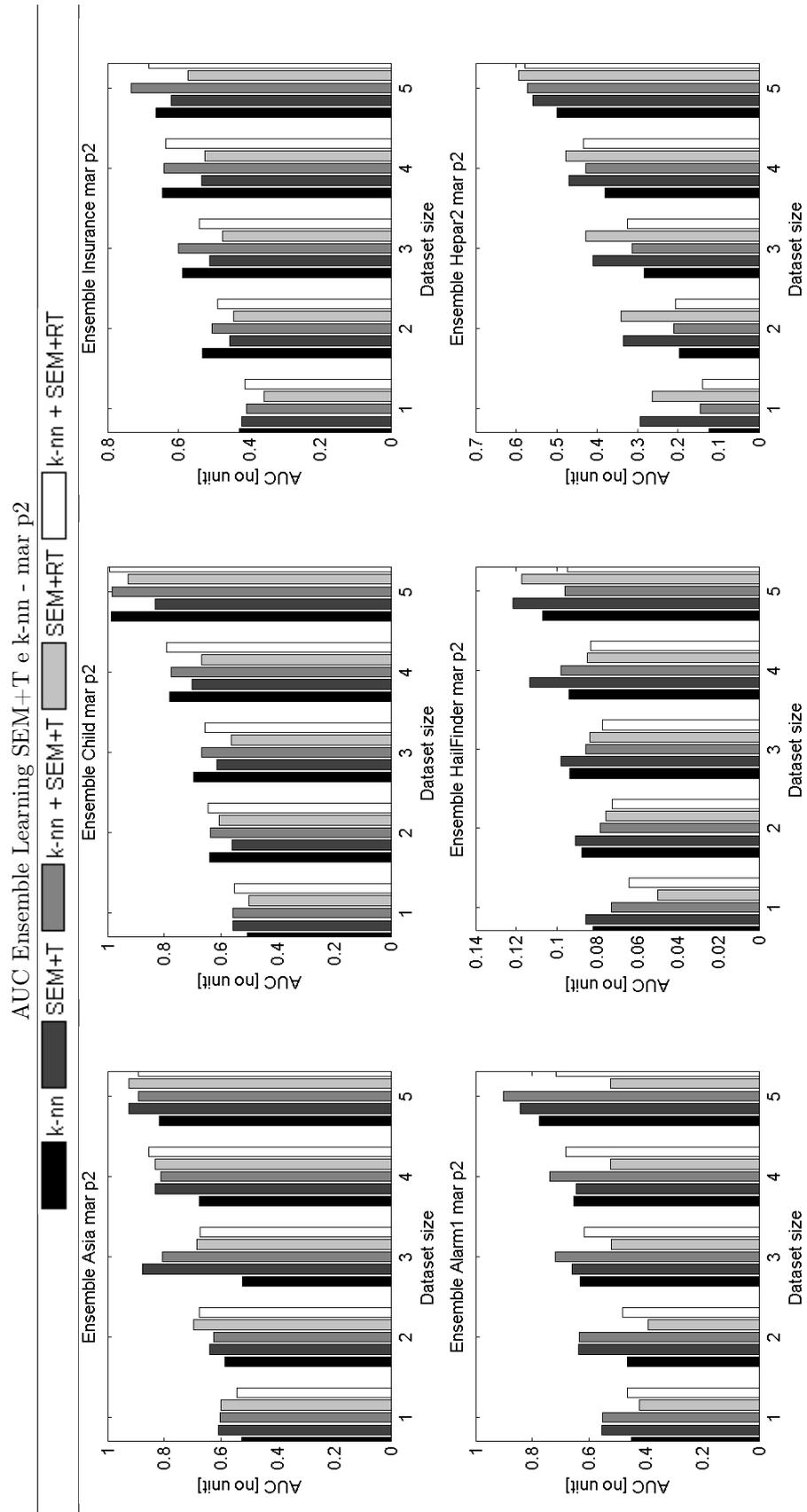


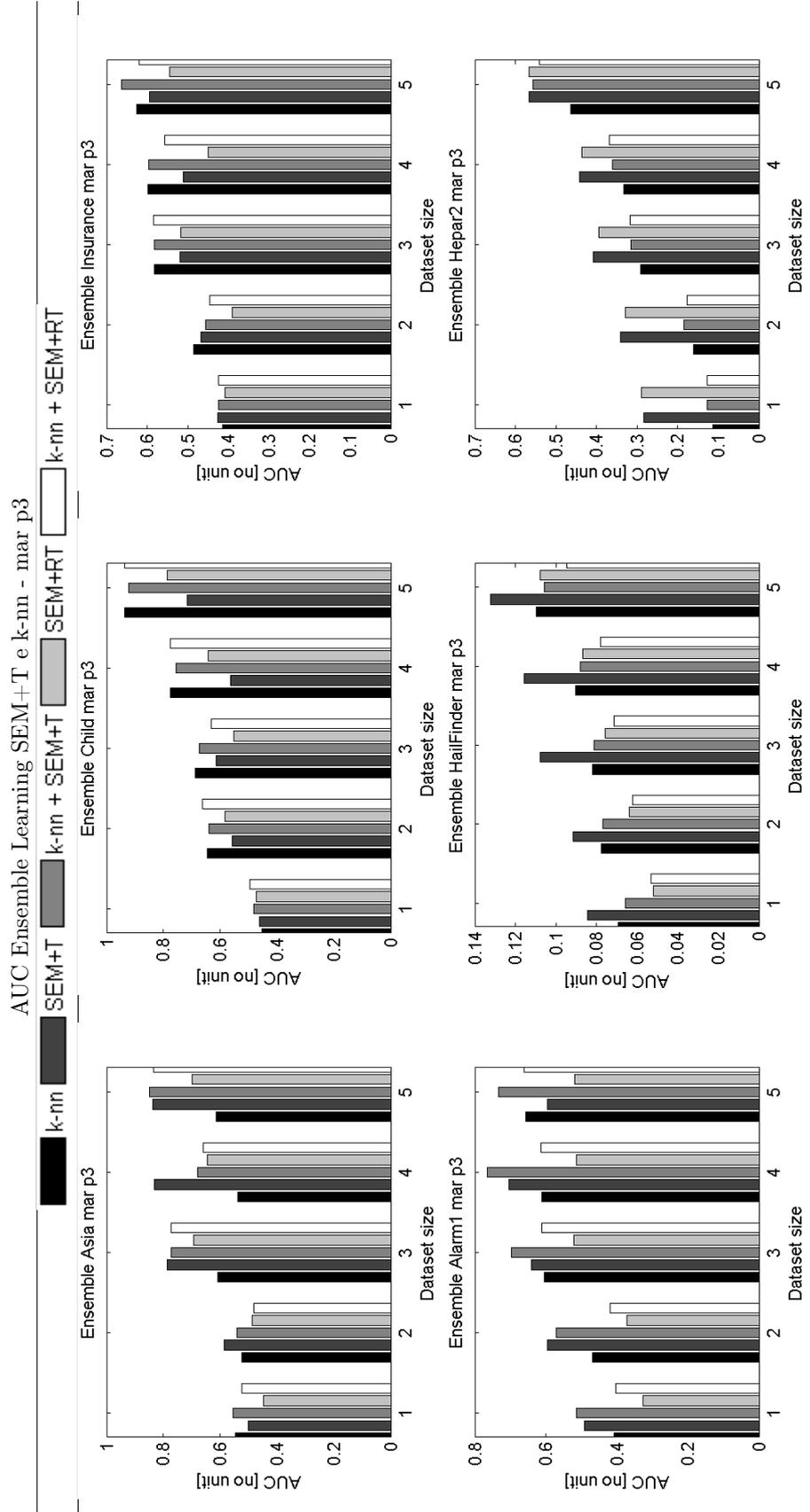
AUC Ensemble Learning SEM+T e k-nn - mcar p2



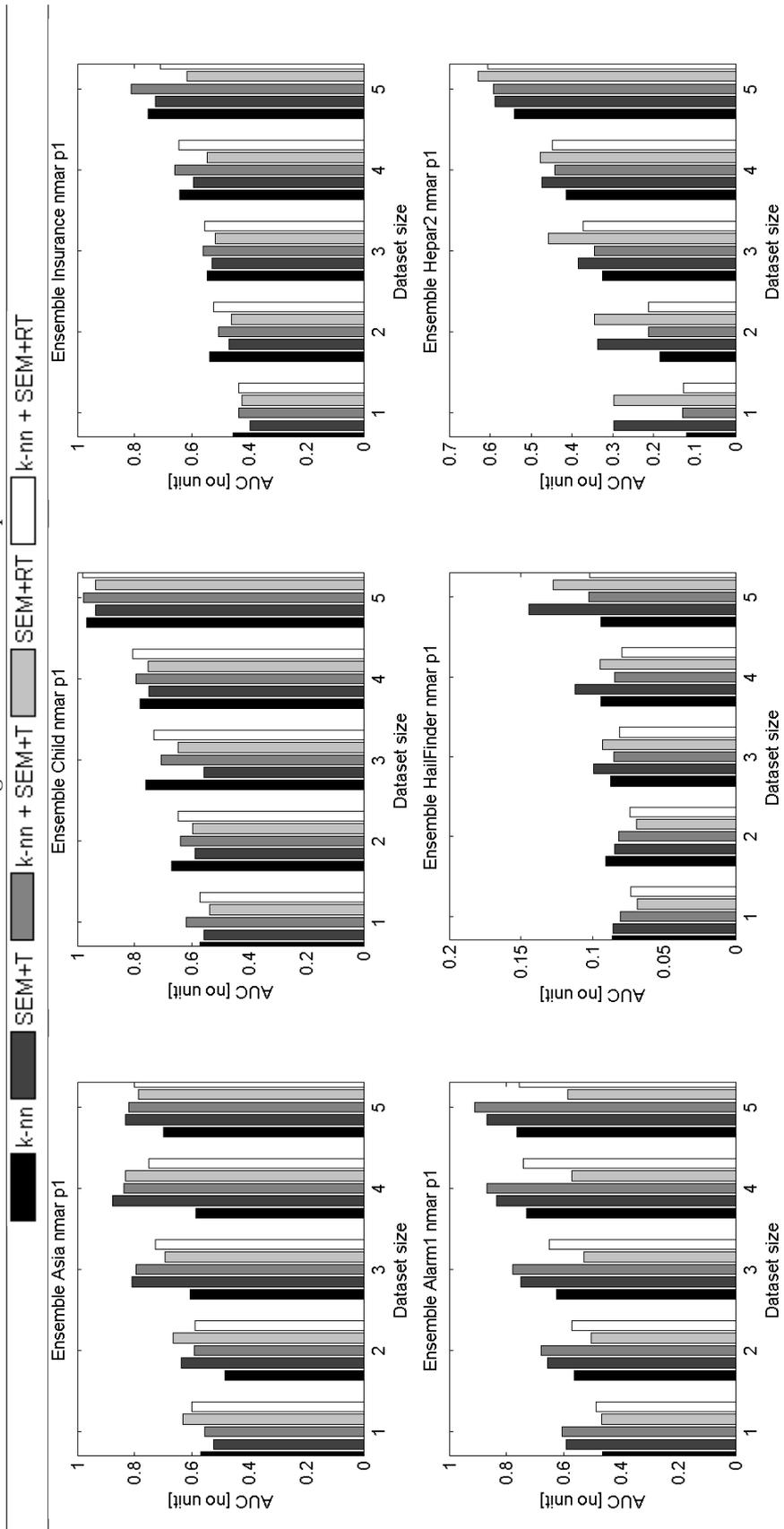




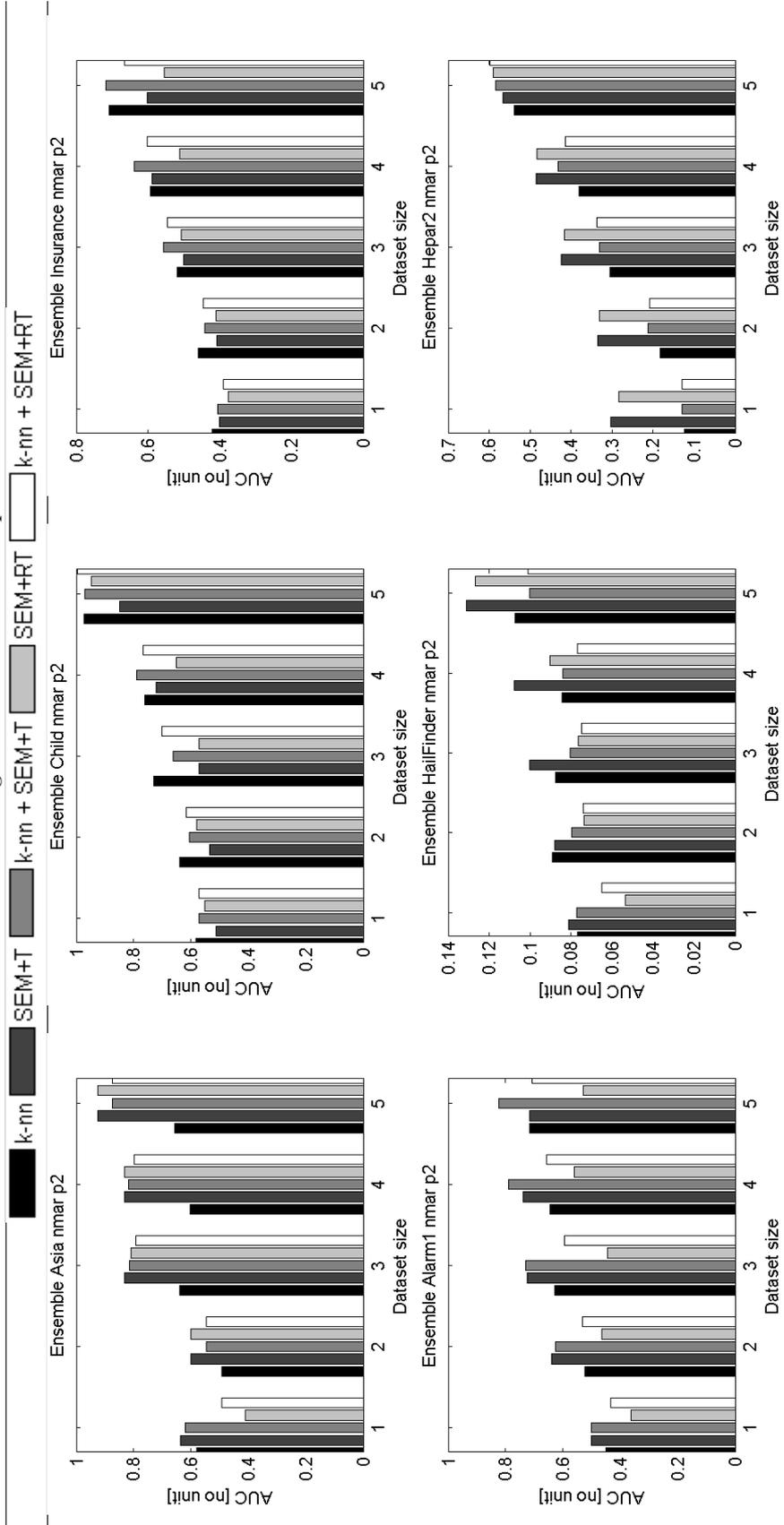


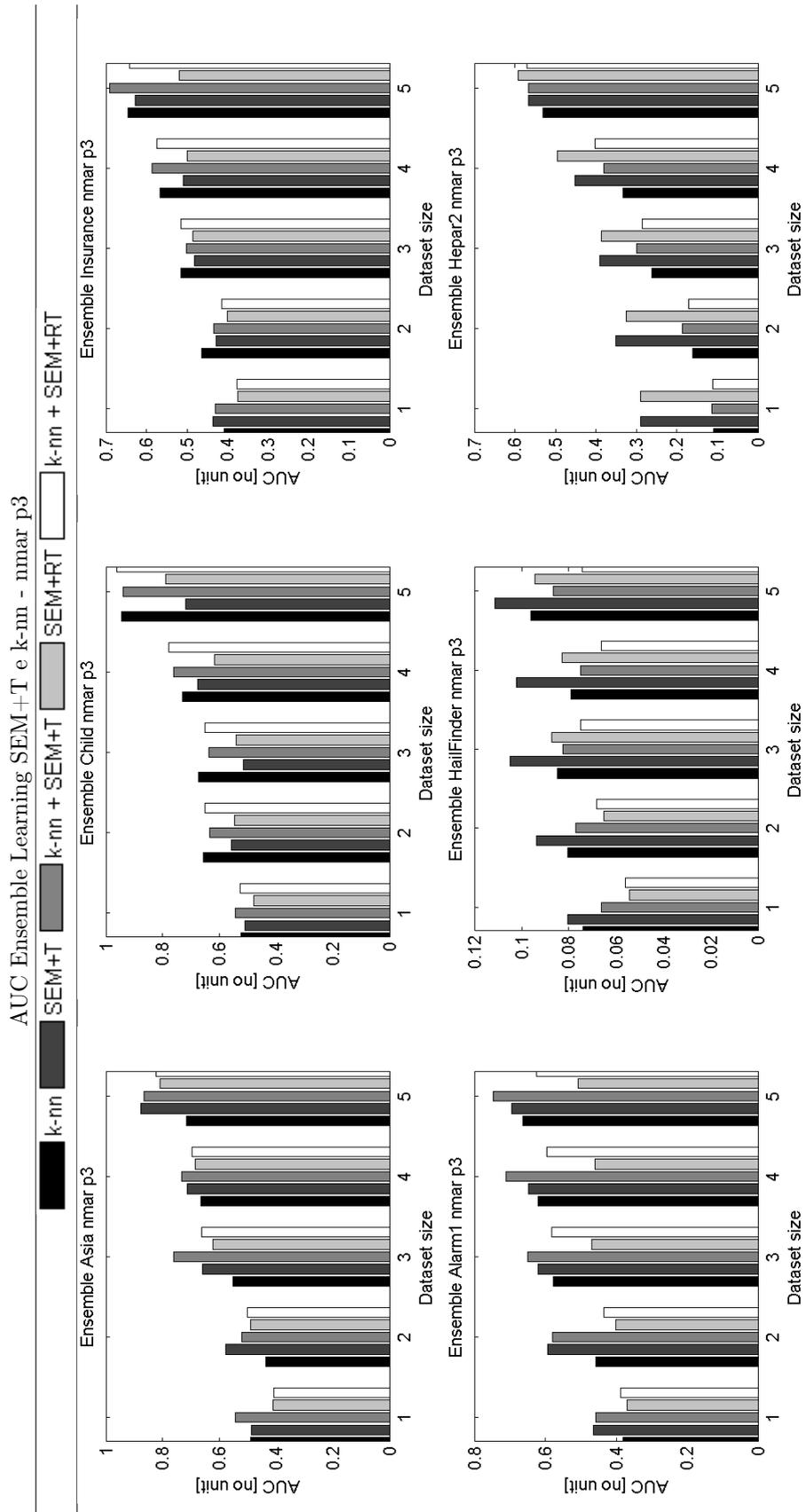


AUC Ensemble Learning SEM+T e k-nn - nmar p1



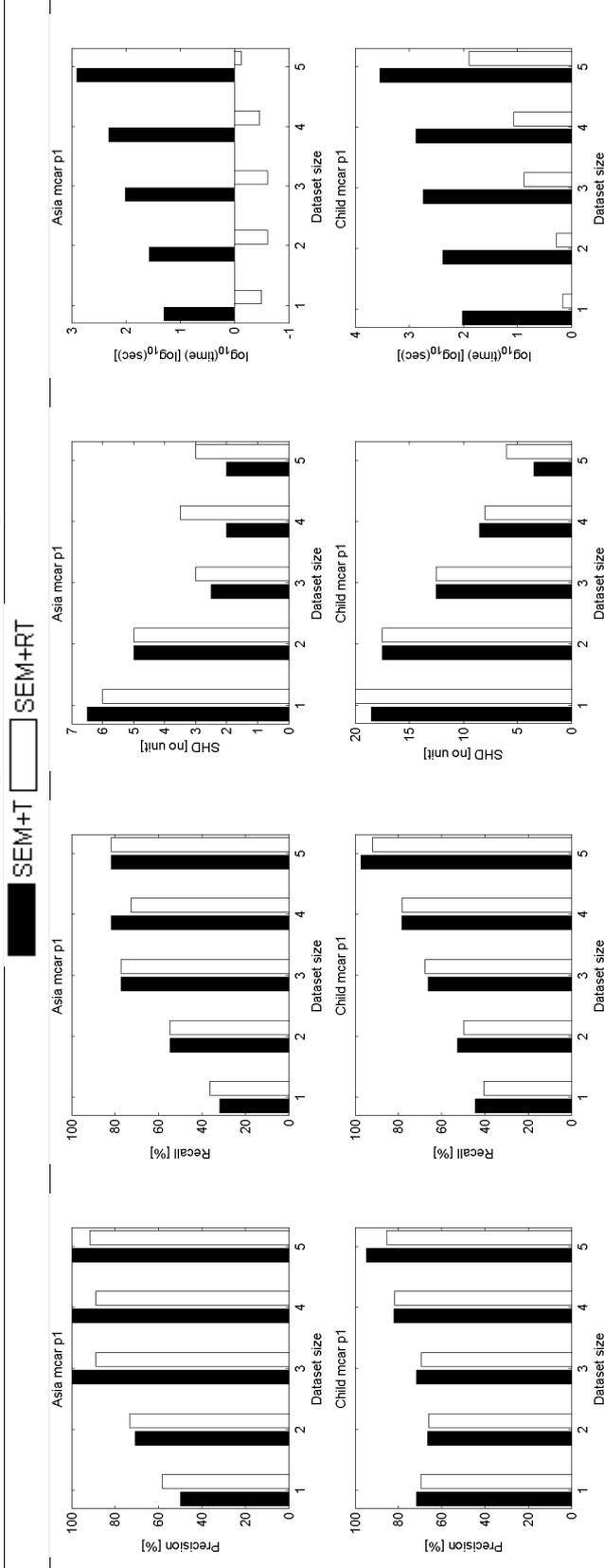
AUC Ensemble Learning SEM+T e k-nn - nmar p2



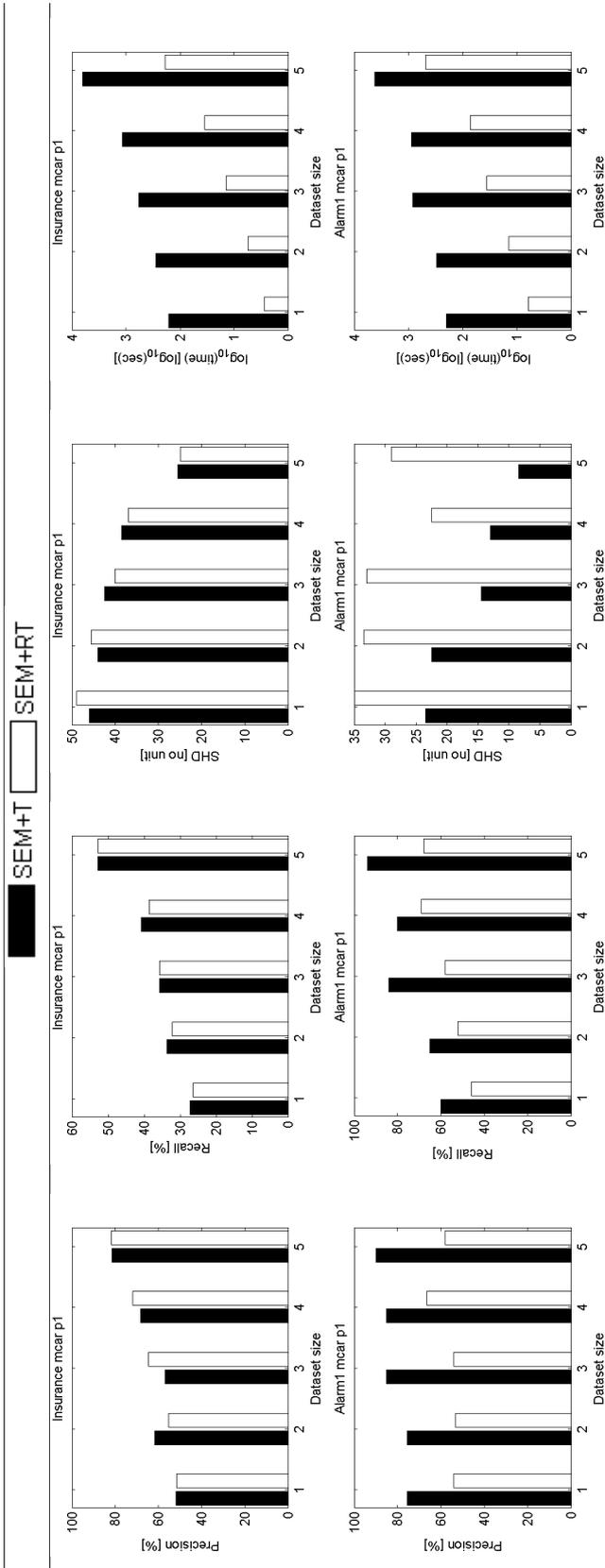


Confronto SEM+T e SEM+RT

Precision-Recall-SHD-Tempo Asia e Child - mcar p1

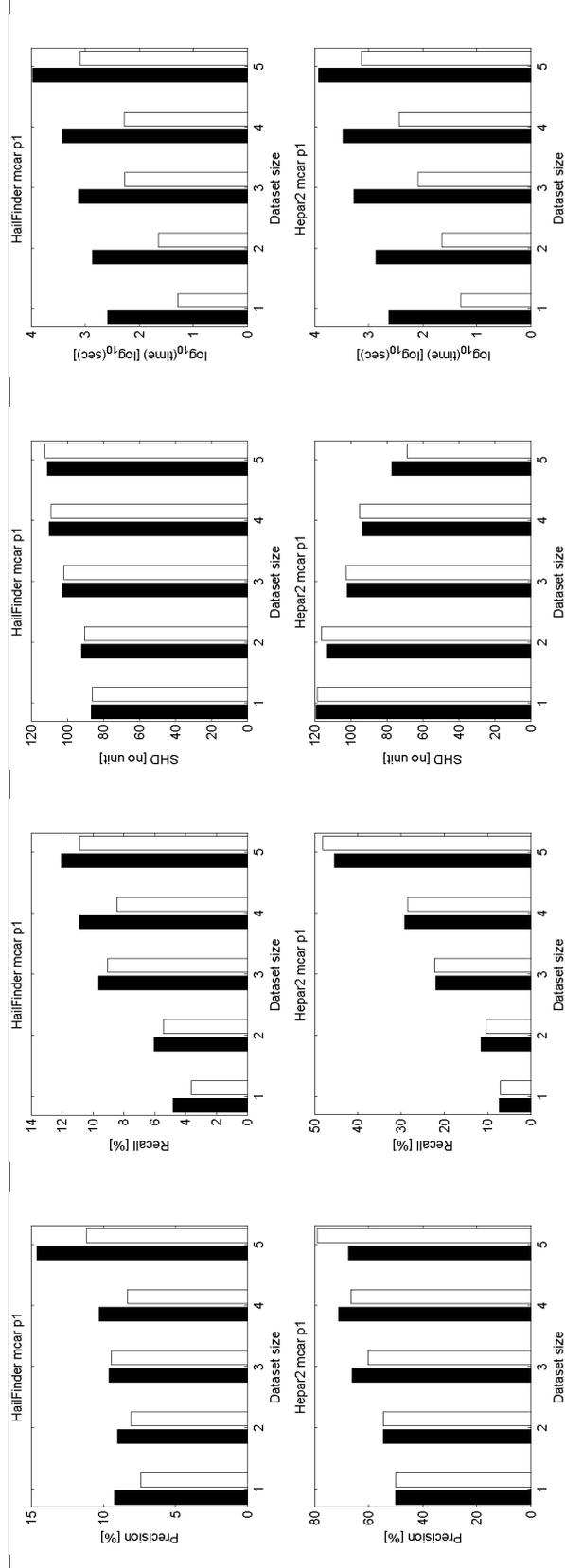


Precision-Recall-SHD-Tempo Insurance e Alarm1 - mcar p1



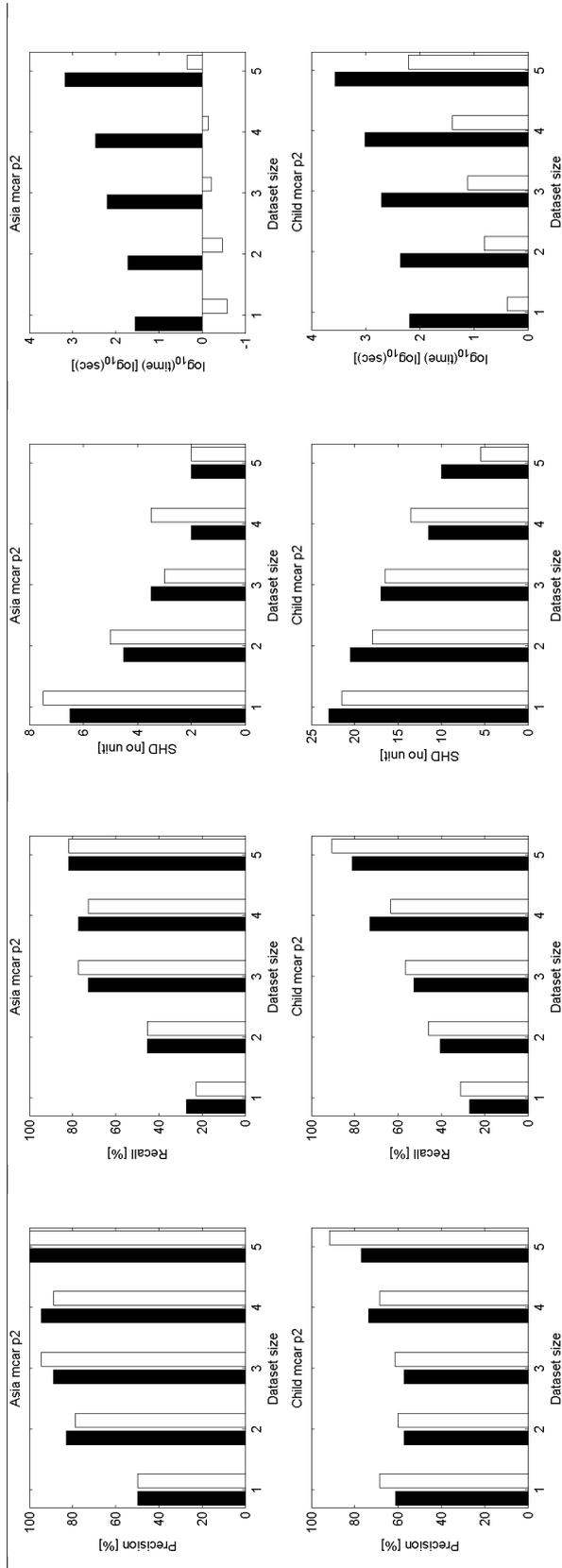
Precision-Recall-SHD-Tempo HailFinder e Hepar2 - mcar p1

SEM+T SEM+RT



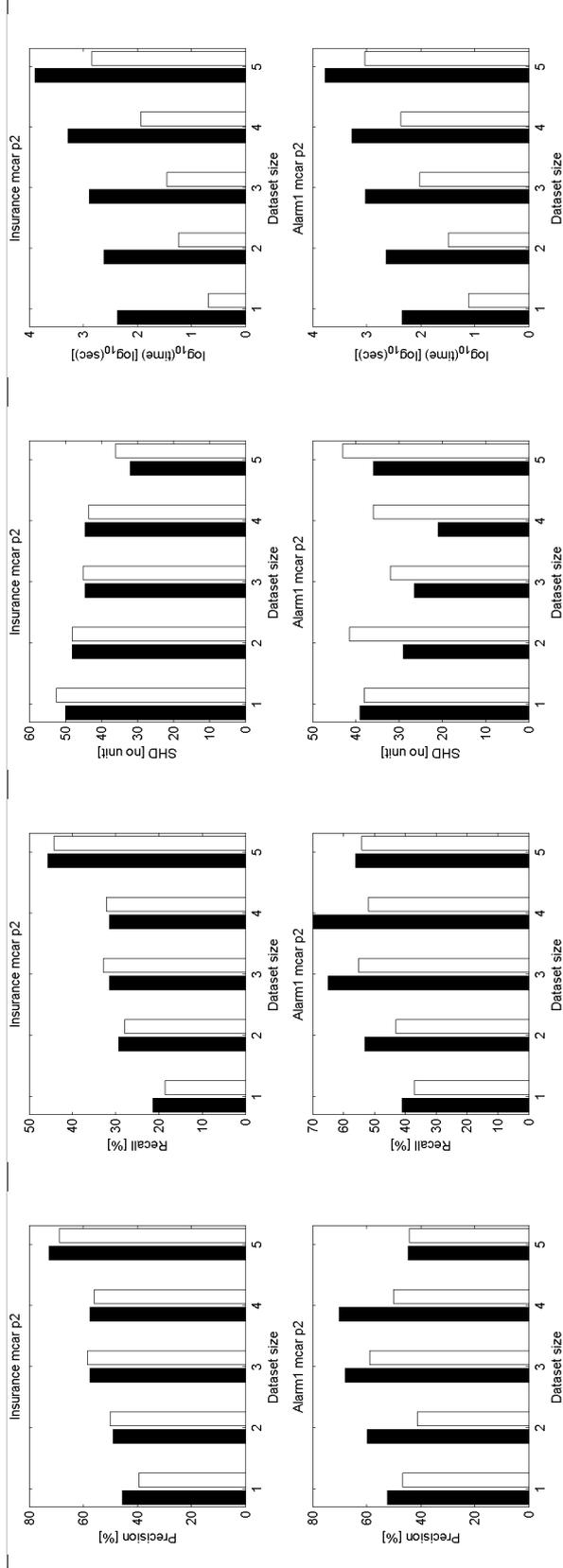
Precision-Recall-SHD-Tempo Asia e Child - mcar p2

SEM+T SEM+RT

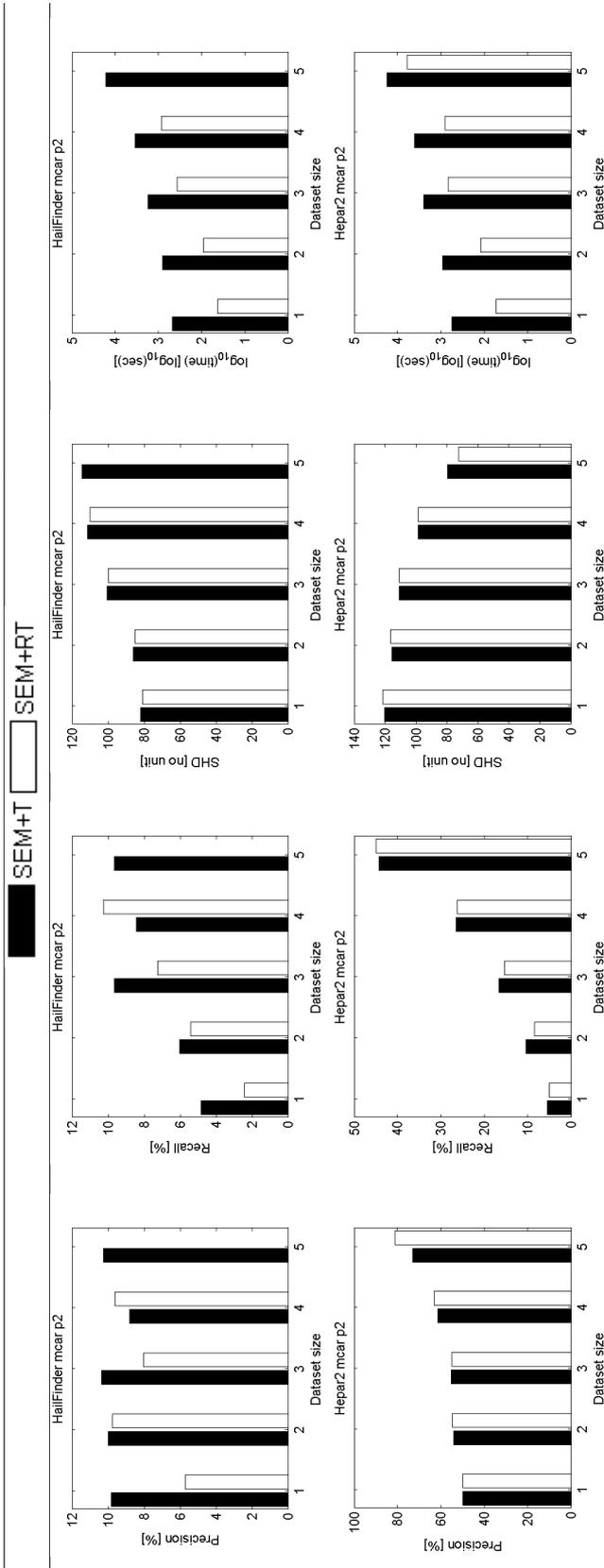


Precision-Recall-SHD-Tempo Insurance e Alarm1 - mcar p2

SEM+T SEM+RT

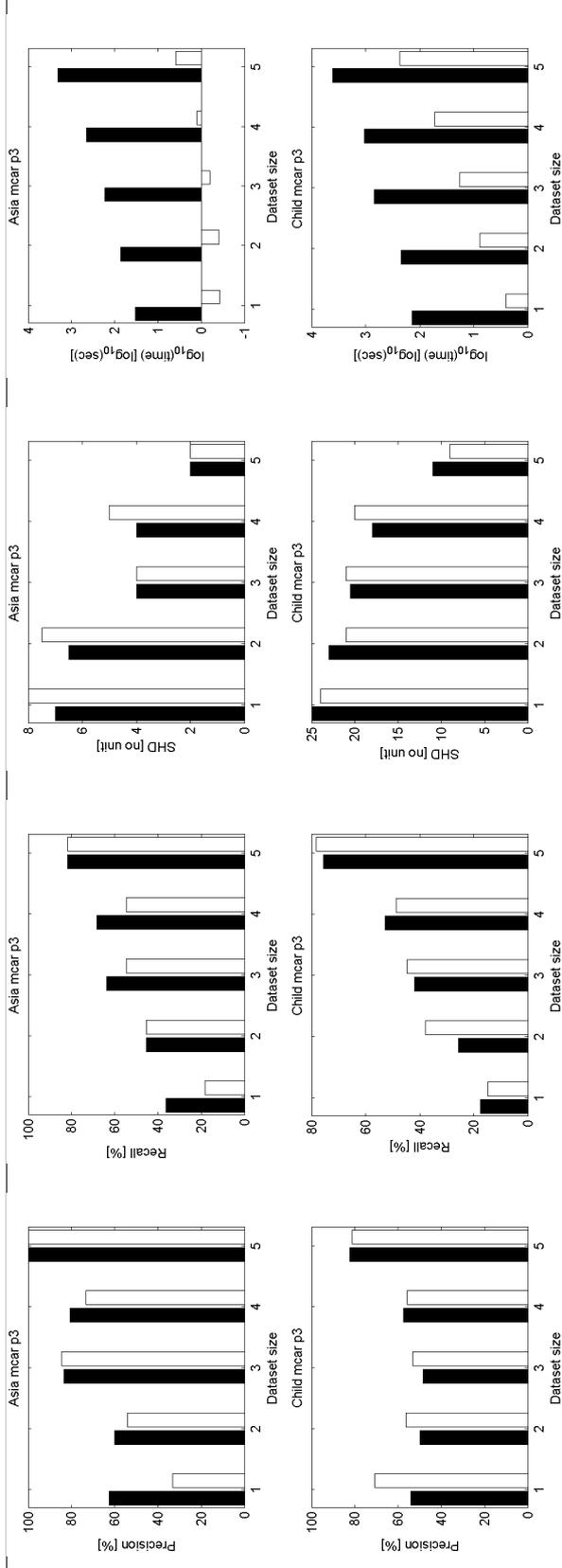


Precision-Recall-SHD-Tempo HailFinder e Hepar2 - mcar p2

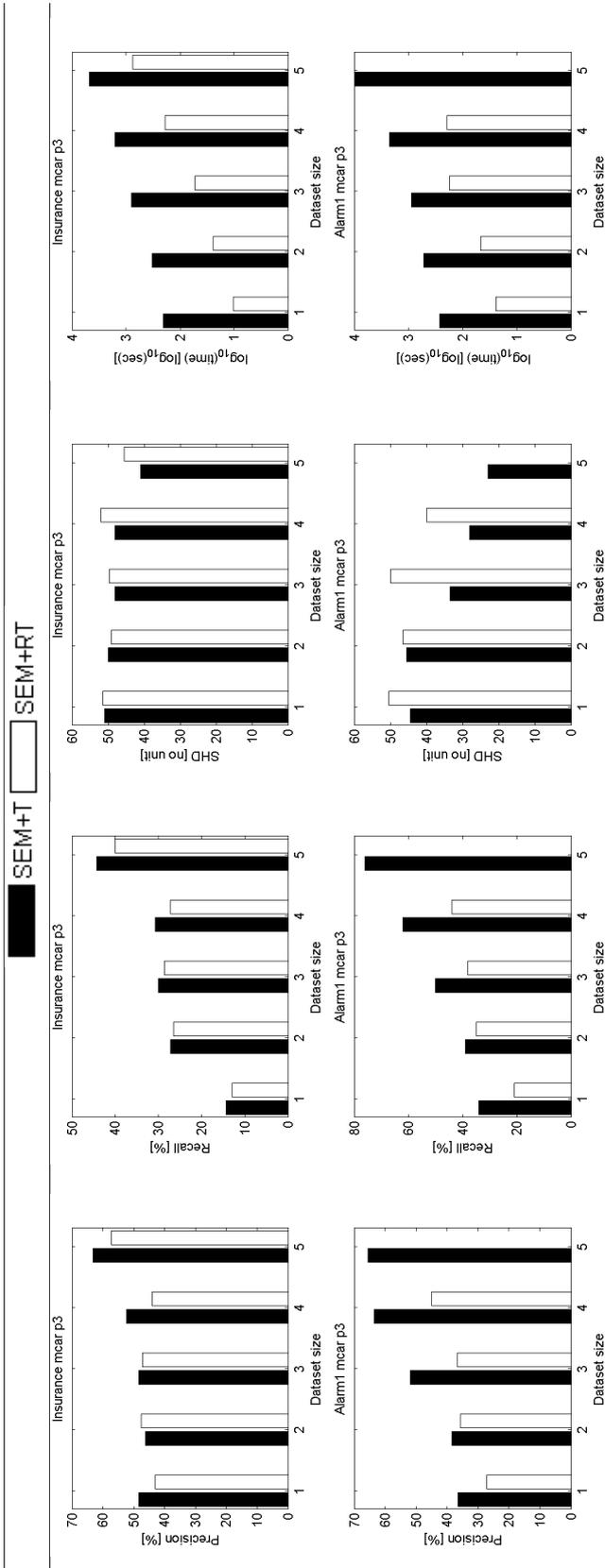


Precision-Recall-SHD-Tempo Asia e Child - mcar p3

SEM+T SEM+RT

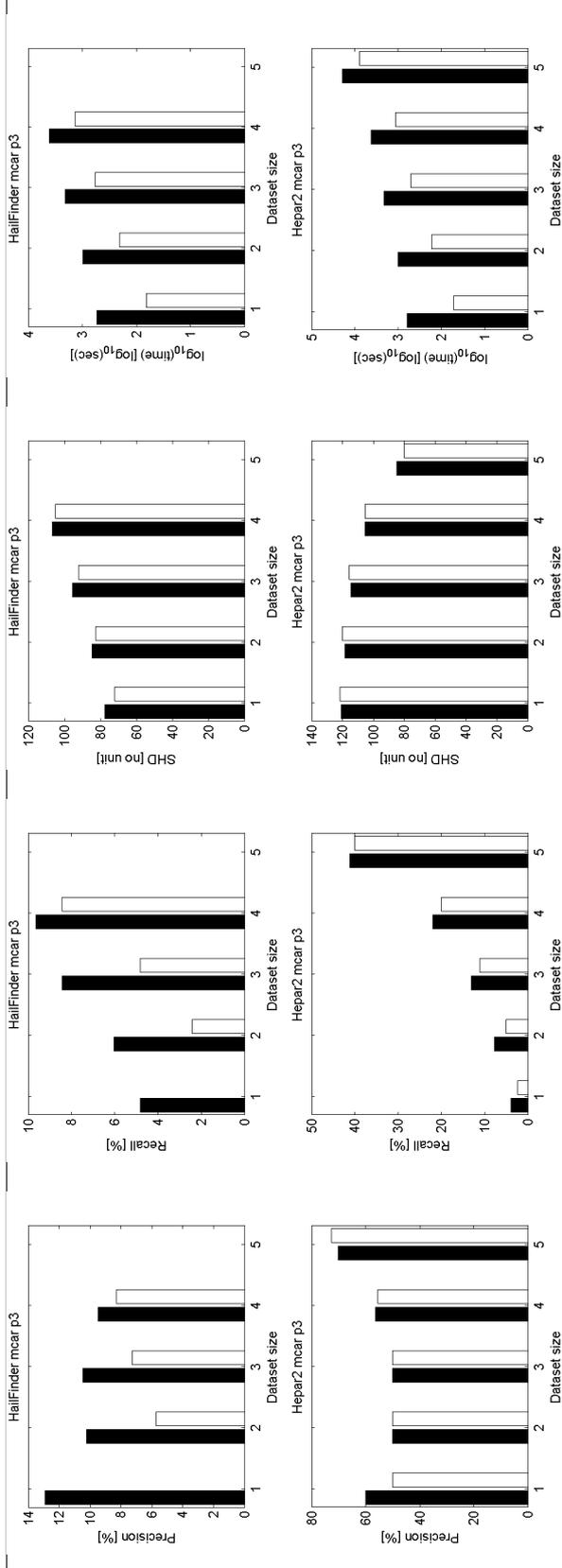


Precision-Recall-SHD-Tempo Insurance e Alarm1 - mcar p3

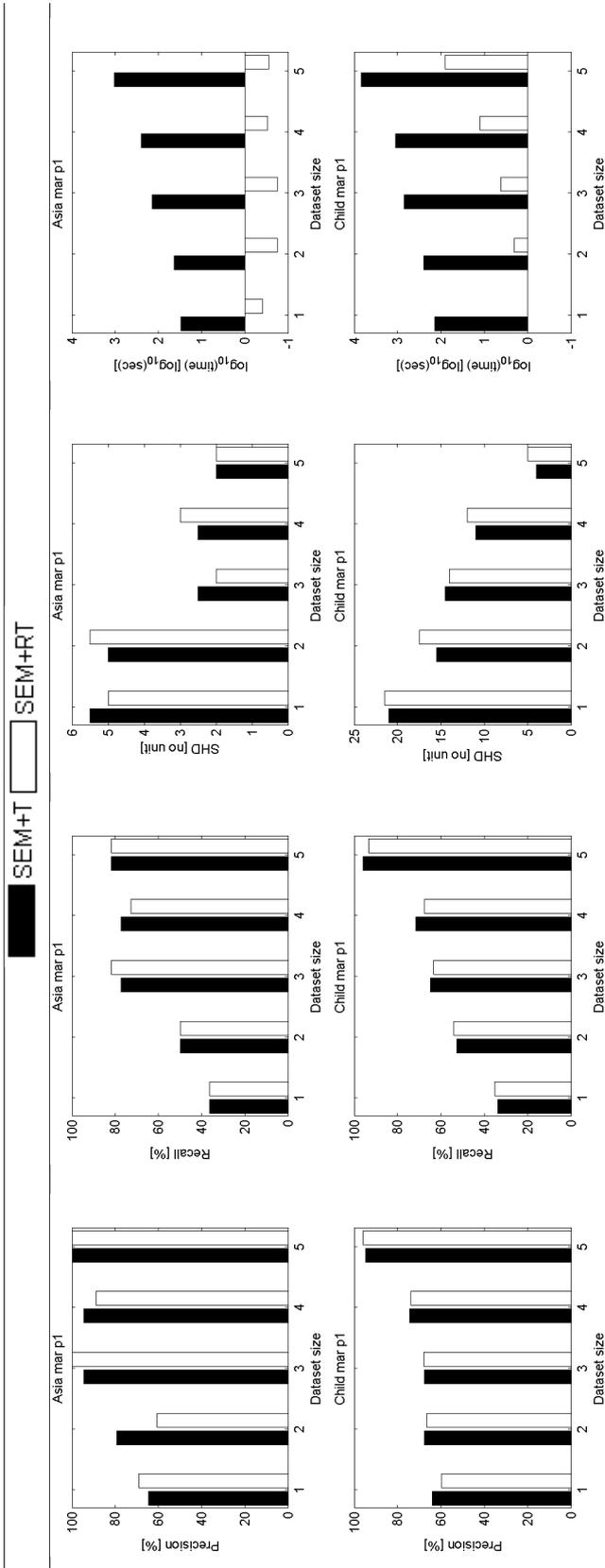


Precision-Recall-SHD-Tempo HailFinder e Hepar2 - mcar p3

SEM+T SEM+RT

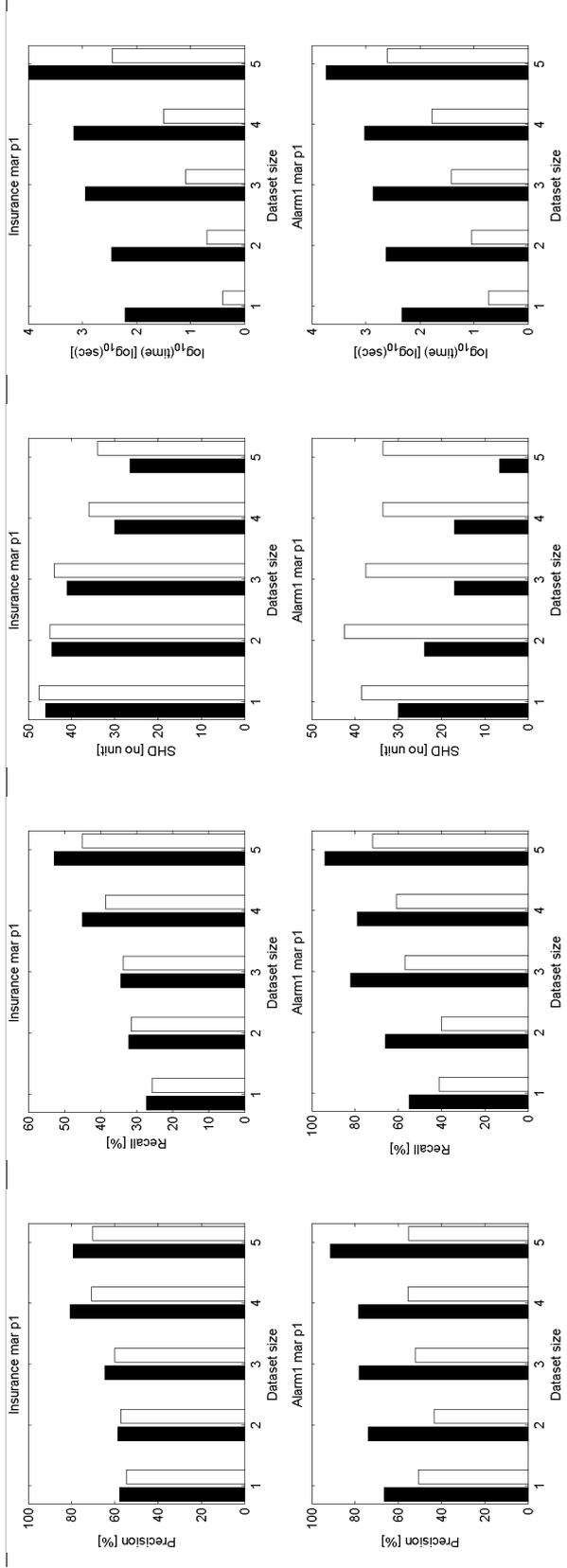


Precision-Recall-SHD-Tempo Asia e Child - mar p1

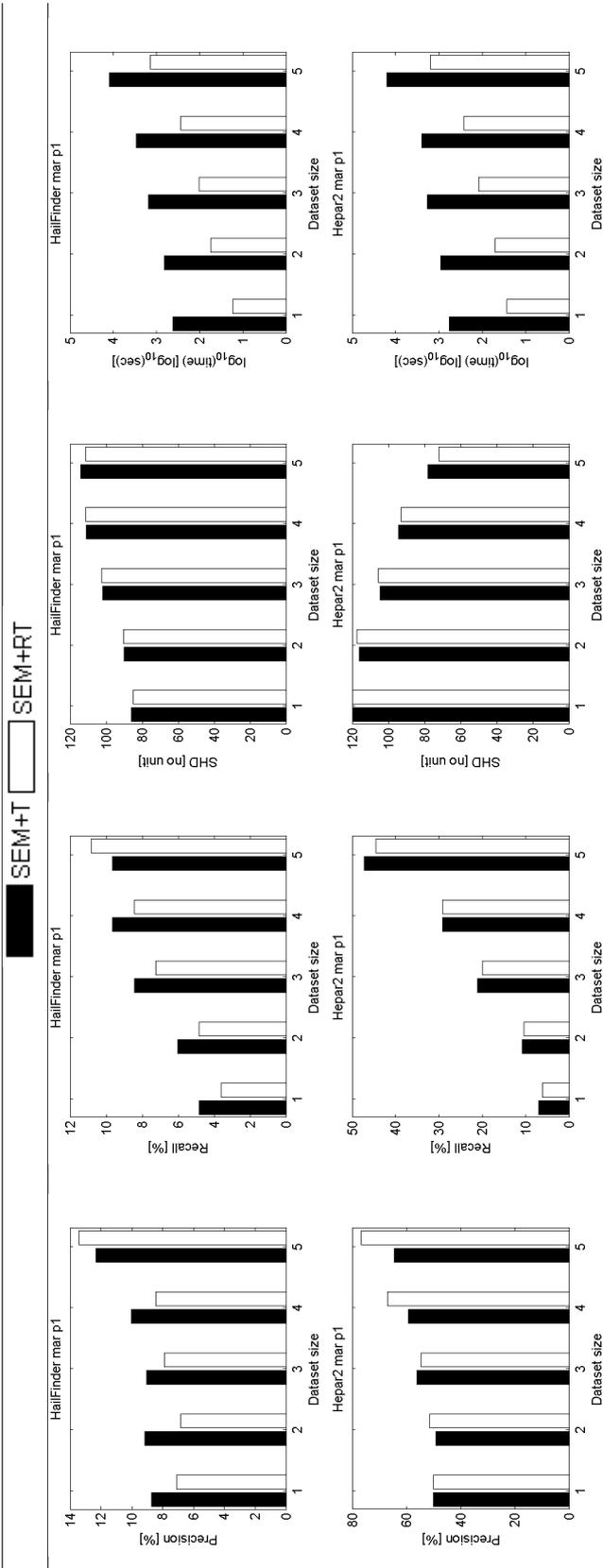


Precision-Recall-SHD-Tempo Insurance e Alarm1 - mar p1

SEM+T SEM+RT

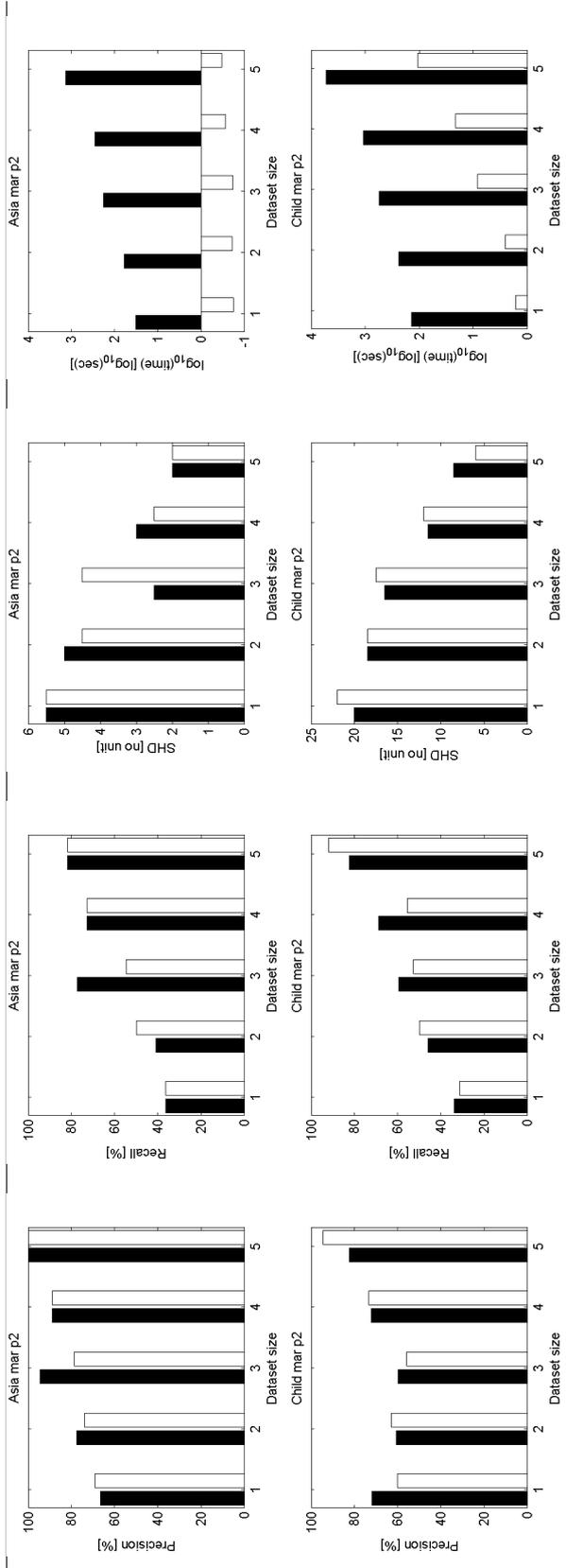


Precision-Recall-SHD-Tempo HailFinder e Hepar2 - mar p1

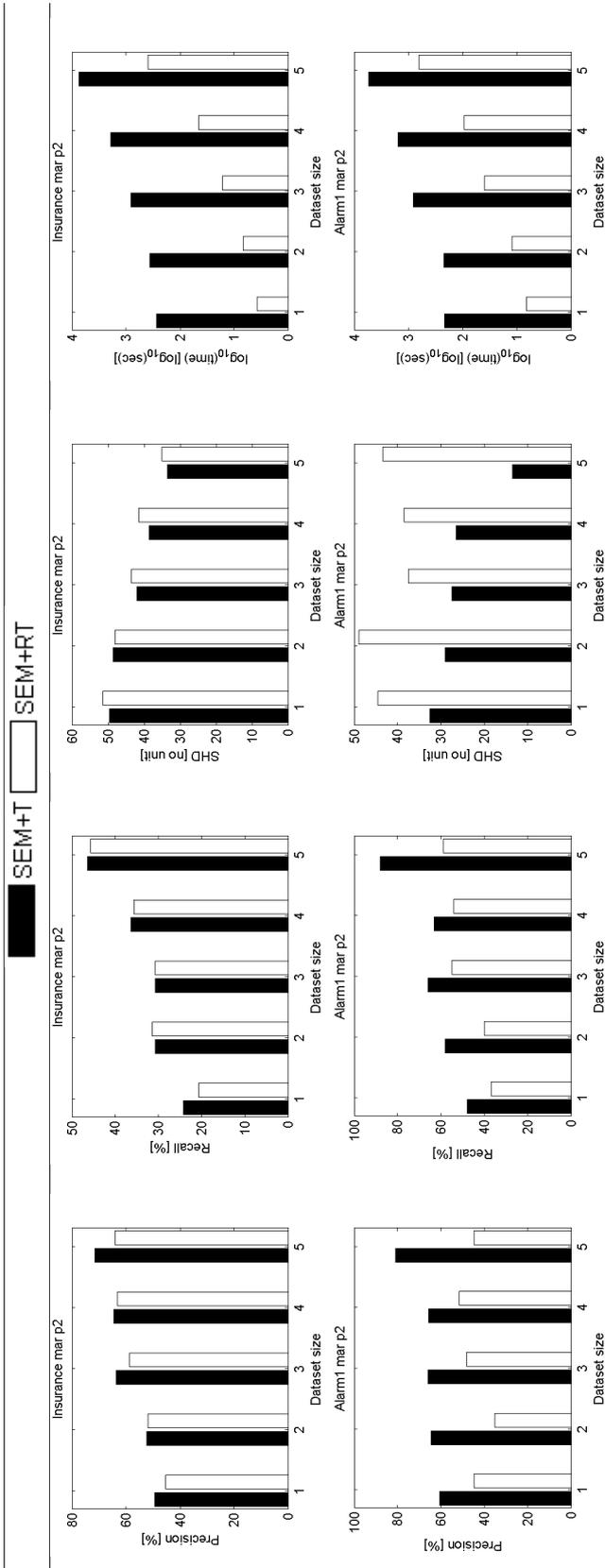


Precision-Recall-SHD-Tempo Asia e Child - mar p2

SEM+T SEM+RT

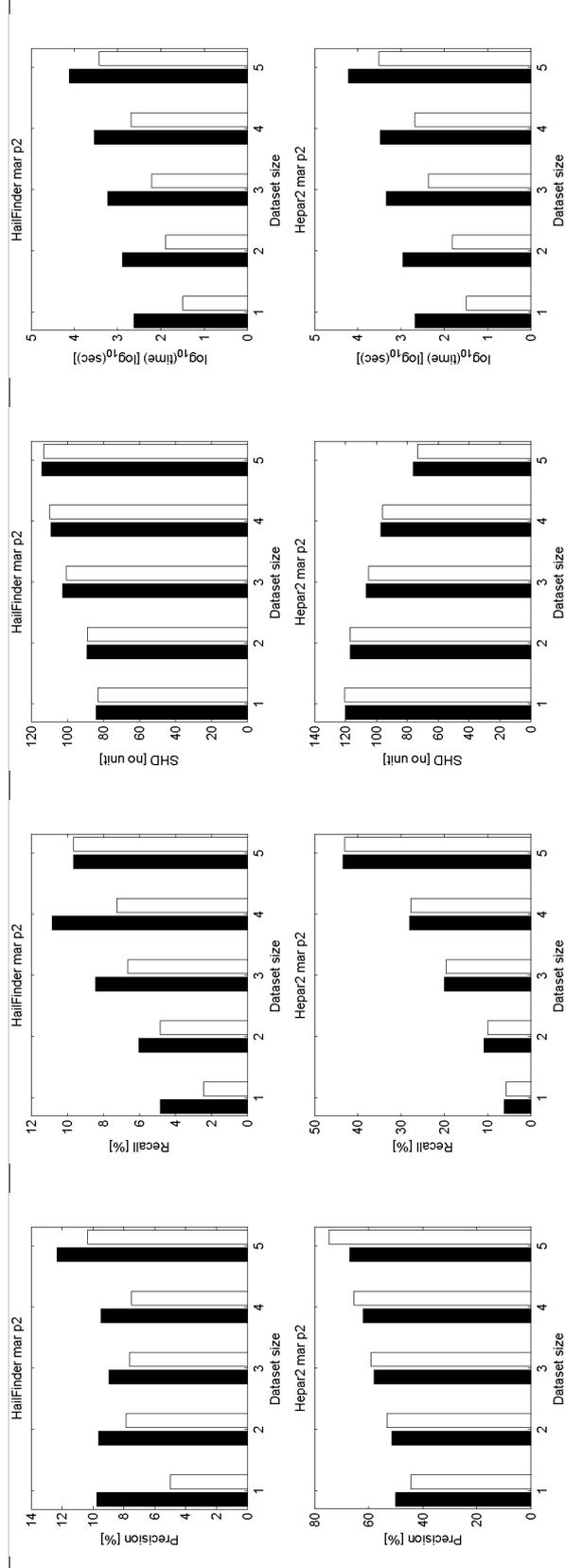


Precision-Recall-SHD-Tempo Insurance e Alarm1 - mar p2

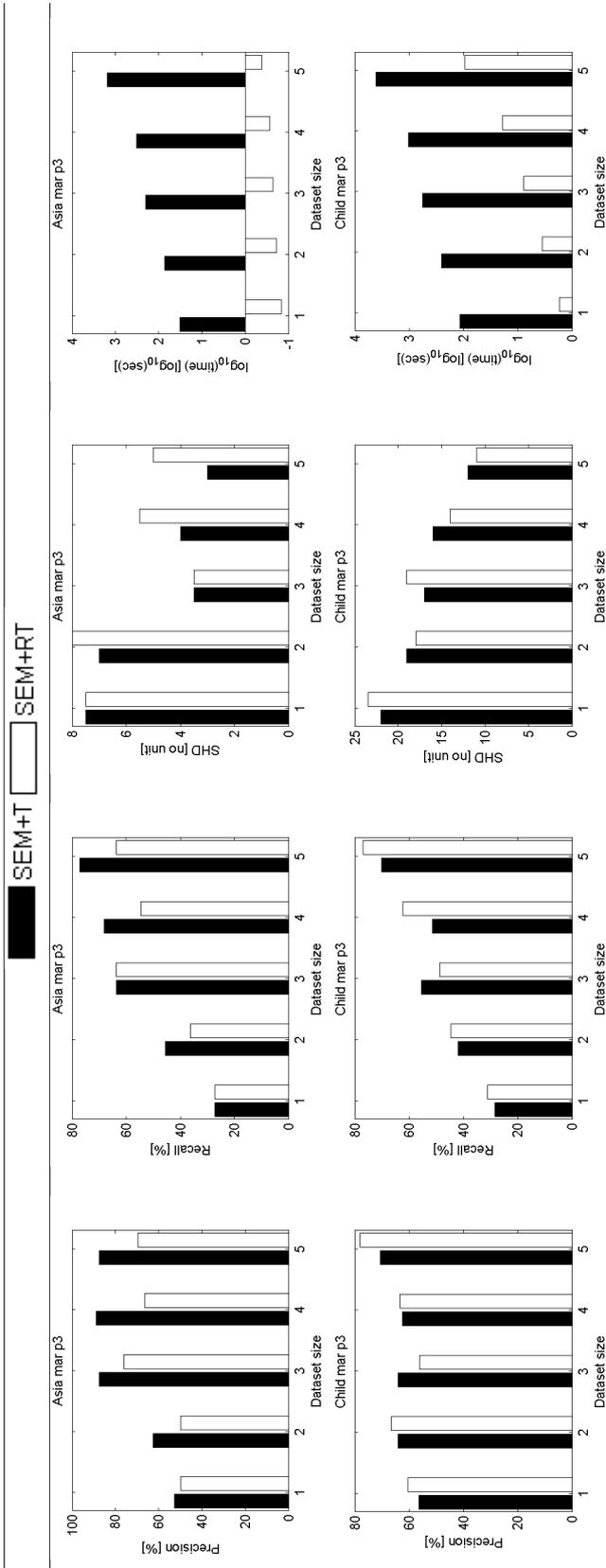


Precision-Recall-SHD-Tempo HailFinder e Hepar2 - mar p2

SEM+T SEM+RT

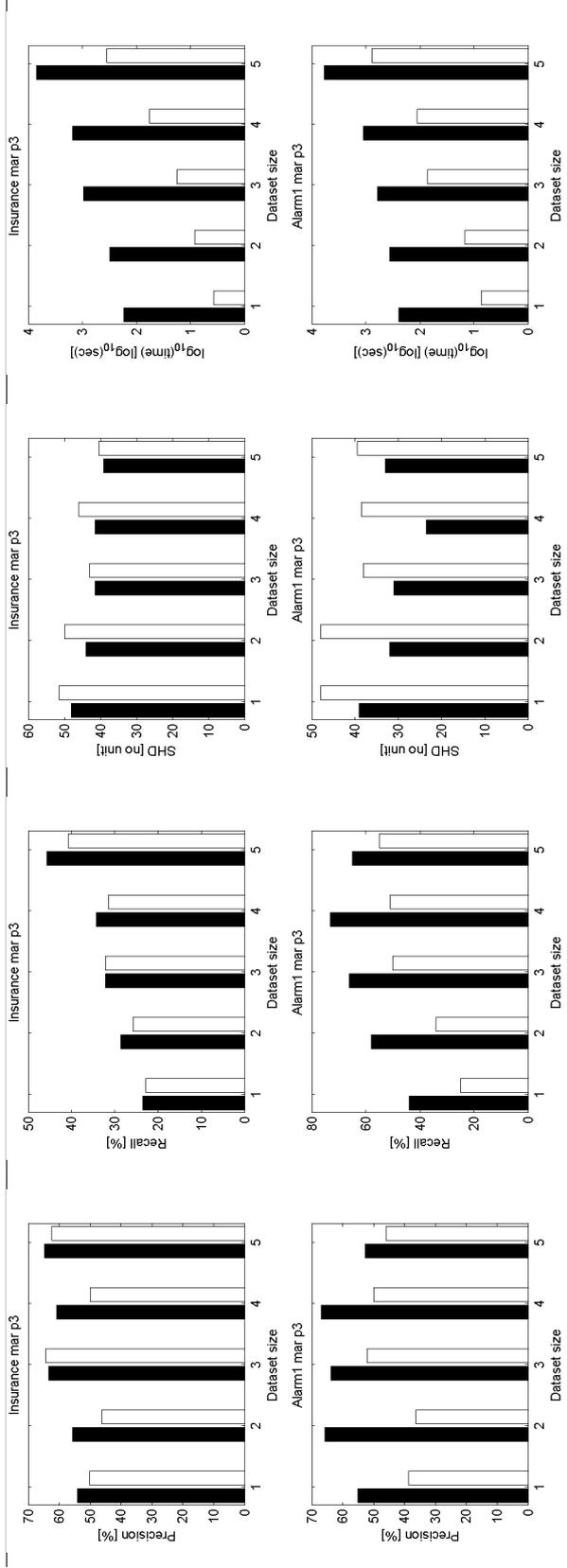


Precision-Recall-SHD-Tempo Asia e Child - mar p3

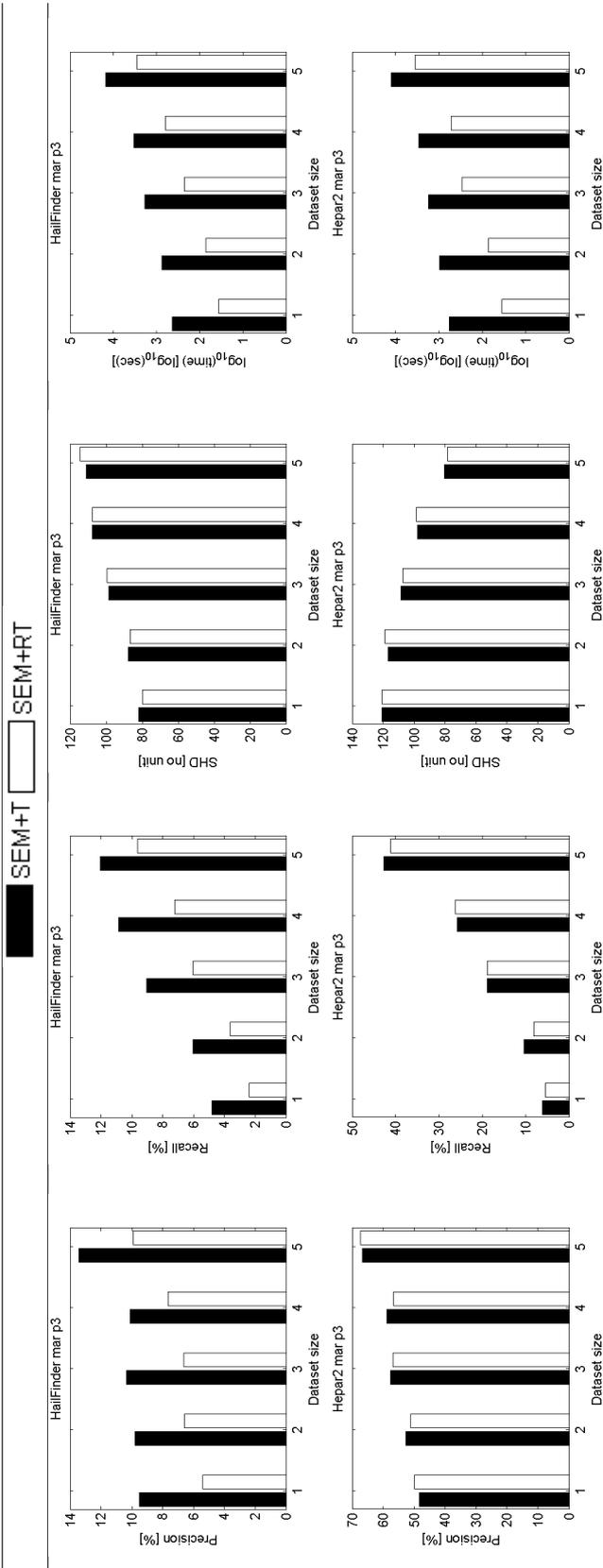


Precision-Recall-SHD-Tempo Insurance e Alarm1 - mar p3

SEM+T SEM+RT

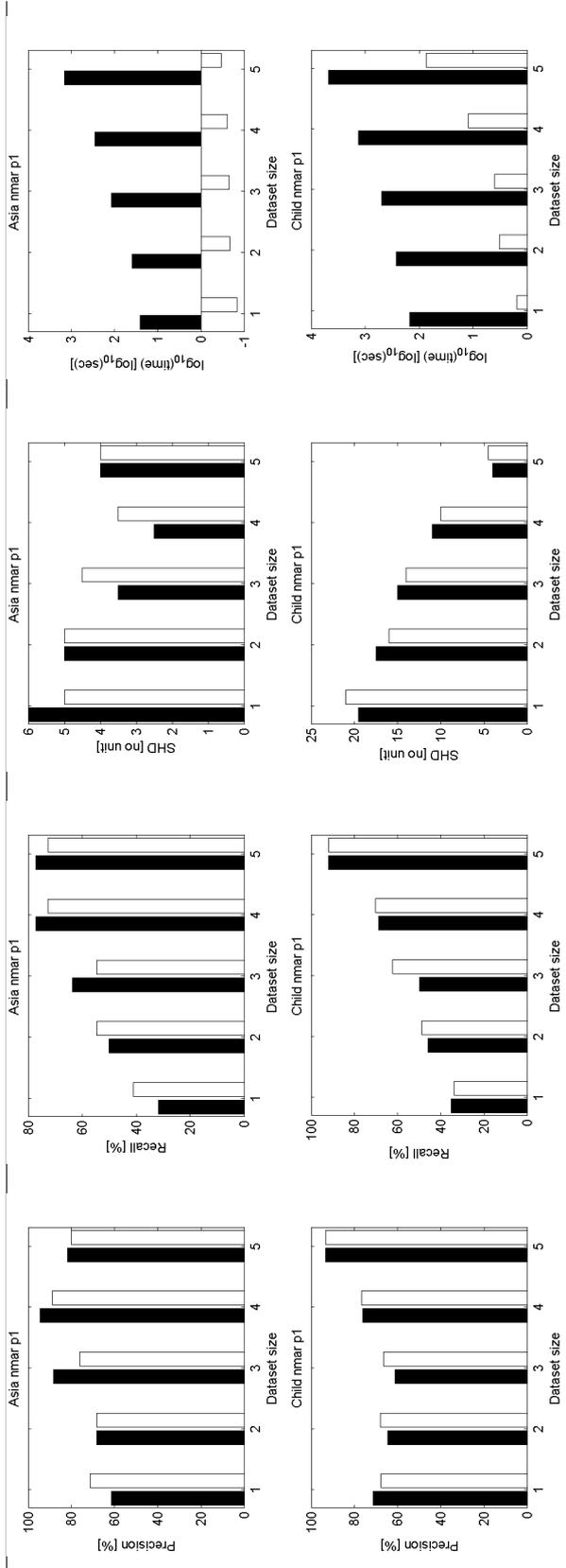


Precision-Recall-SHD-Tempo HailFinder e Hepar2 - mar p3

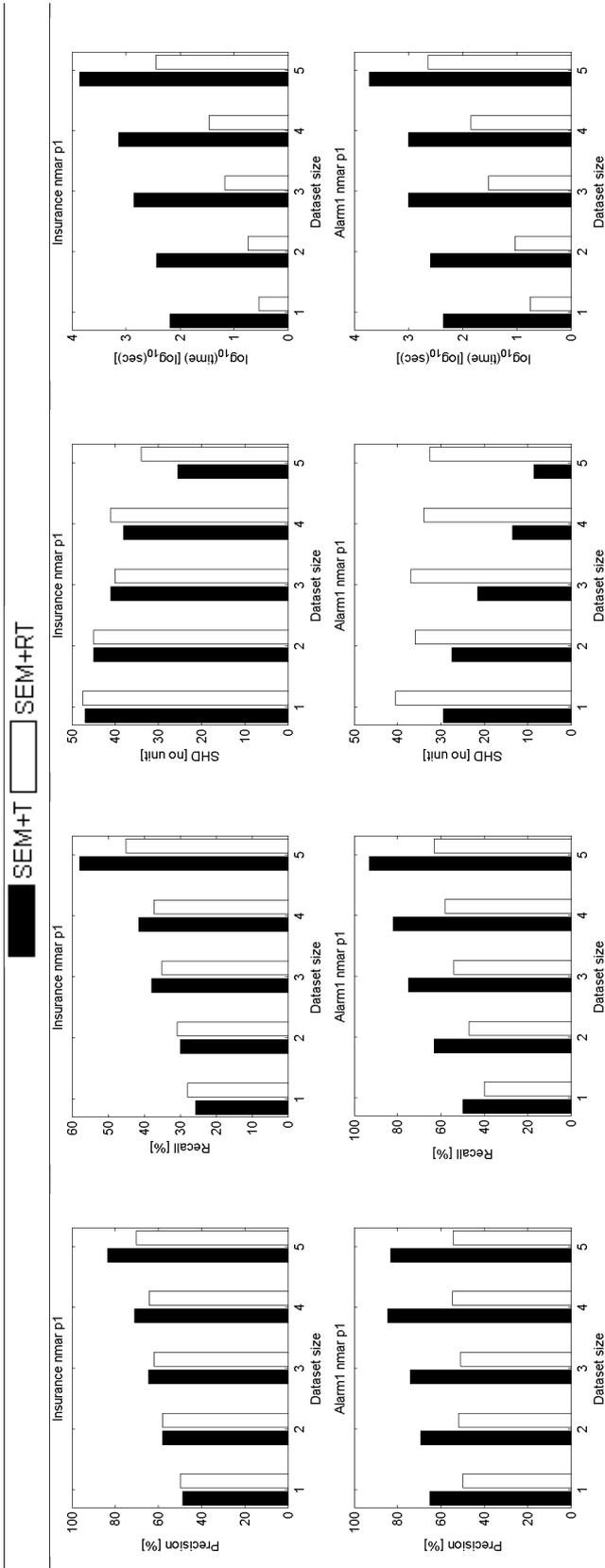


Precision-Recall-SHD-Tempo Asia e Child - nmar p1

SEM+T SEM+RT

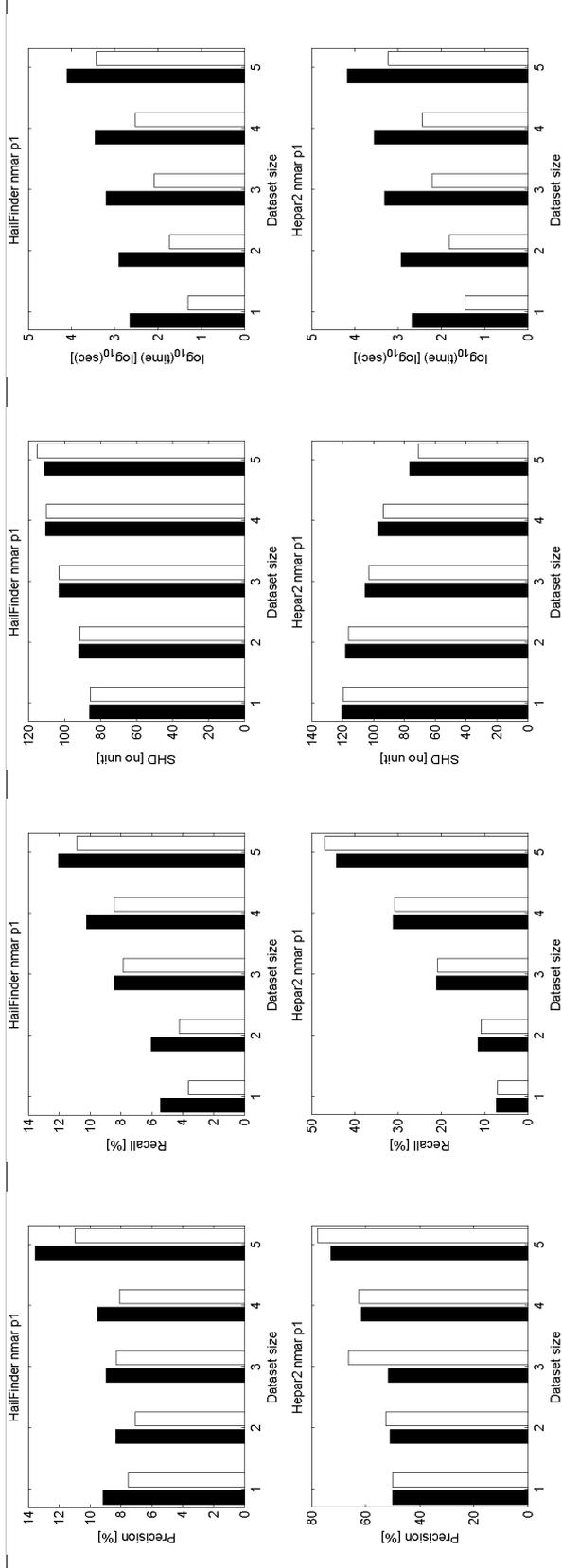


Precision-Recall-SHD-Tempo Insurance e Alarm1 - nmar p1



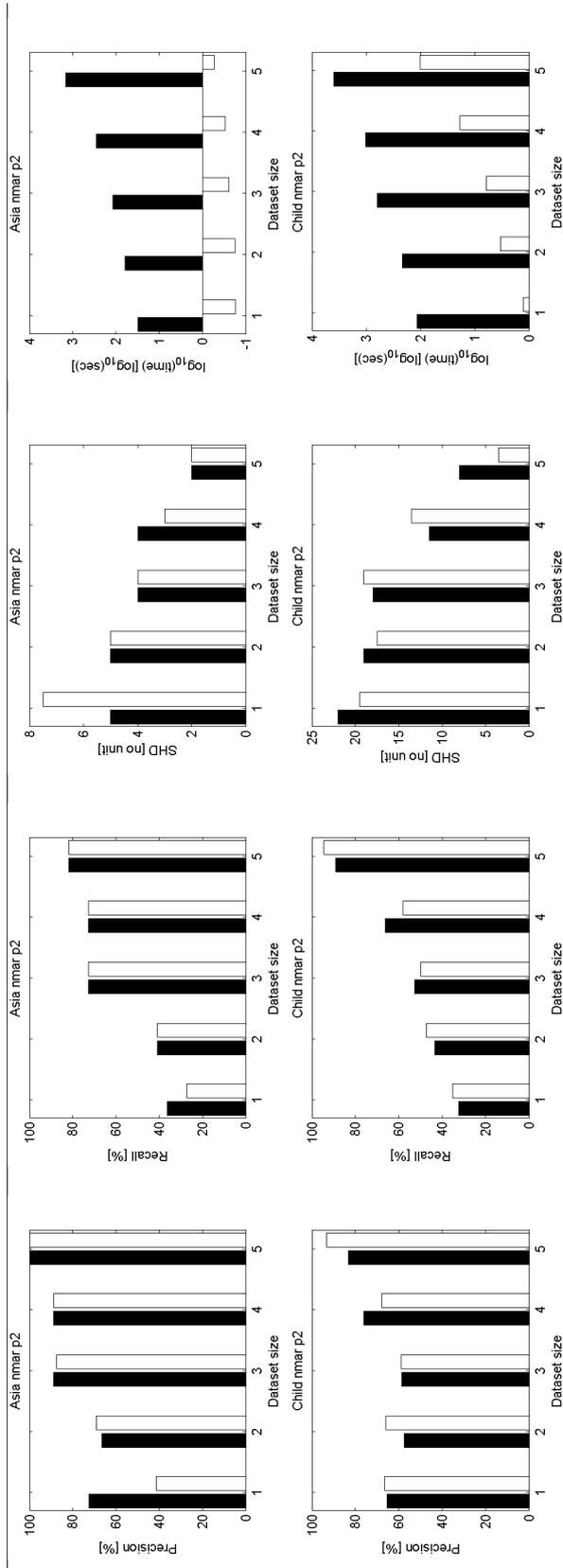
Precision-Recall-SHD-Tempo HaiFinder e Hepar2 - nmar p1

SEM+T SEM+RT



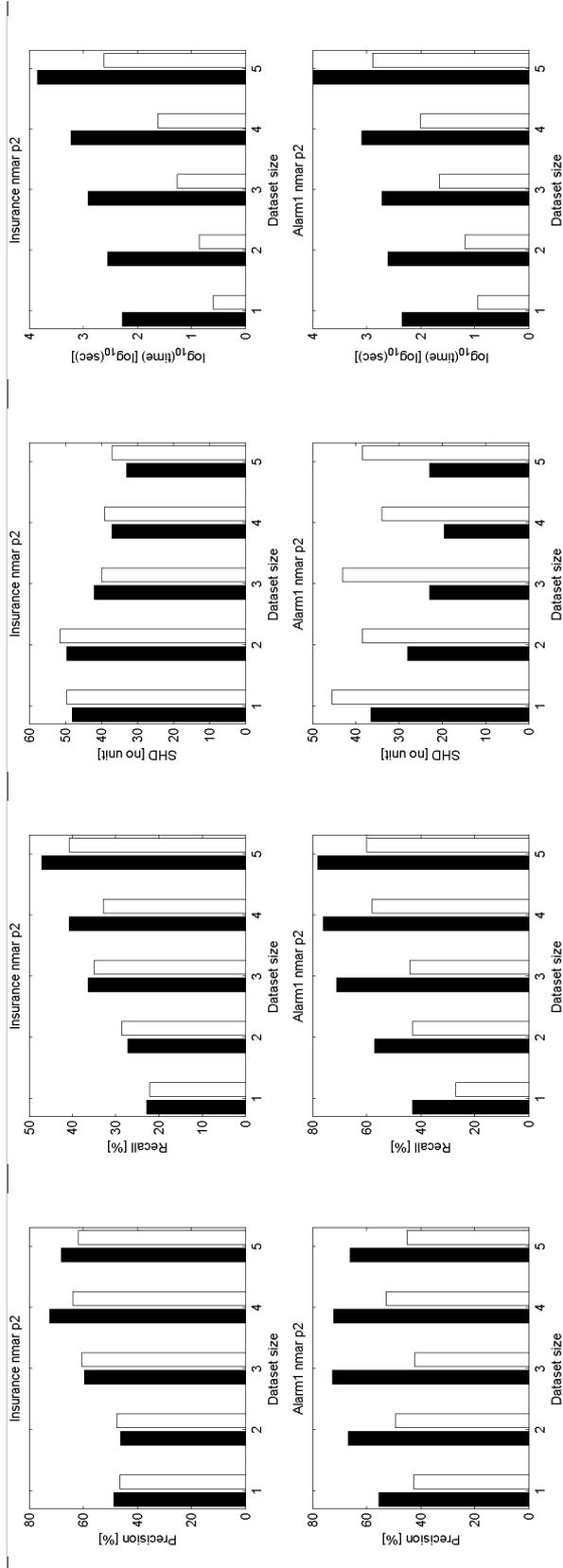
Precision-Recall-SHD-Tempo Asia e Child - nmar p2

SEM+T SEM+RT

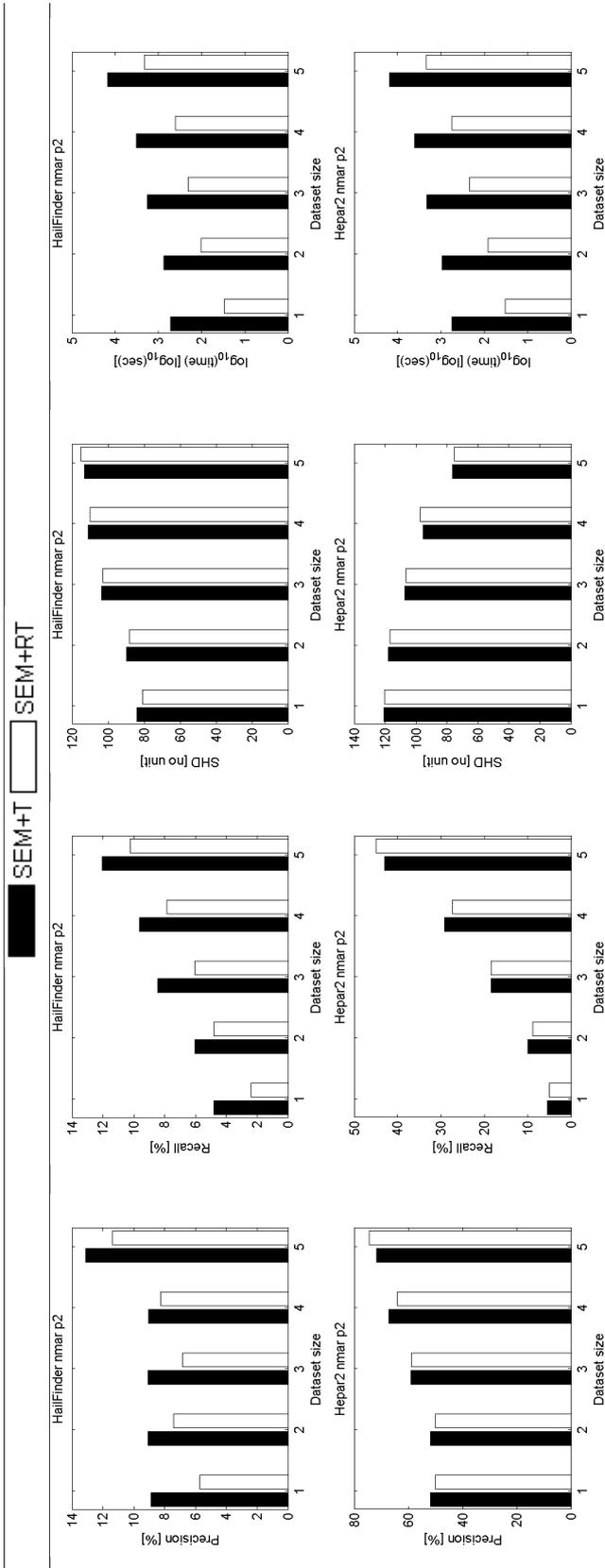


Precision-Recall-SHD-Tempo Insurance e Alarm1 - nmar p2

SEM+T SEM+RT

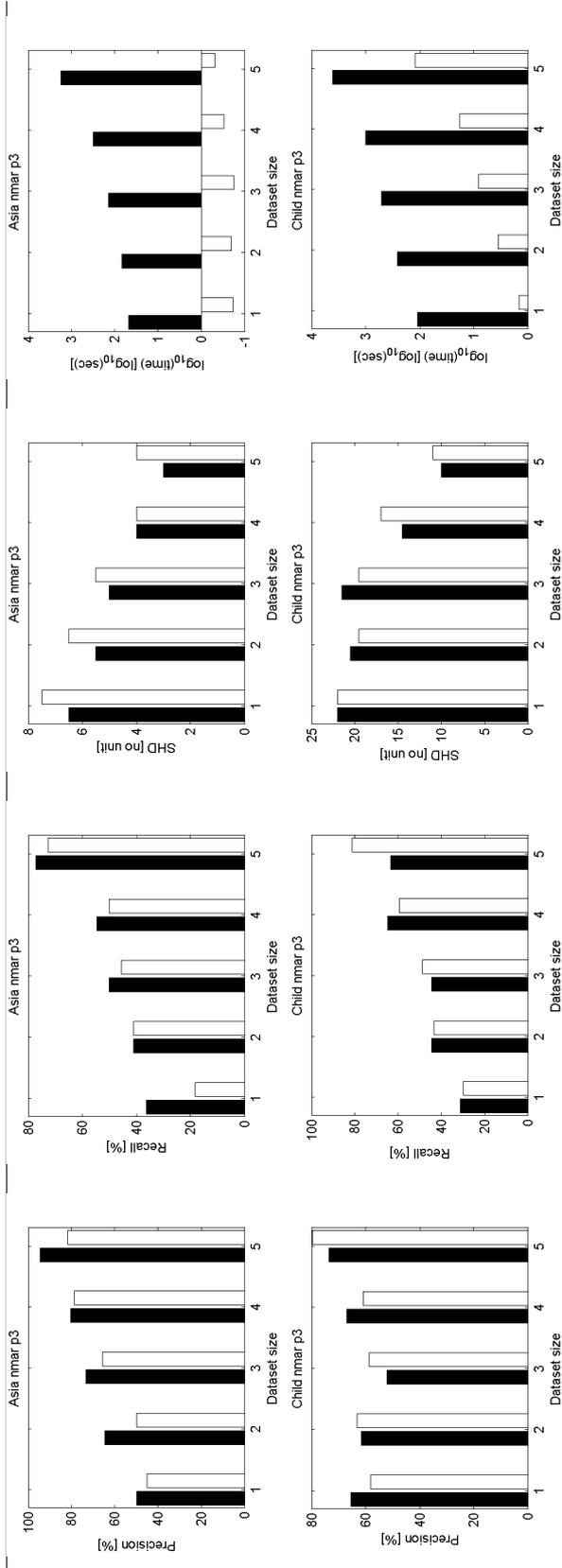


Precision-Recall-SHD-Tempo HailFinder e Hepar2 - nmar p2

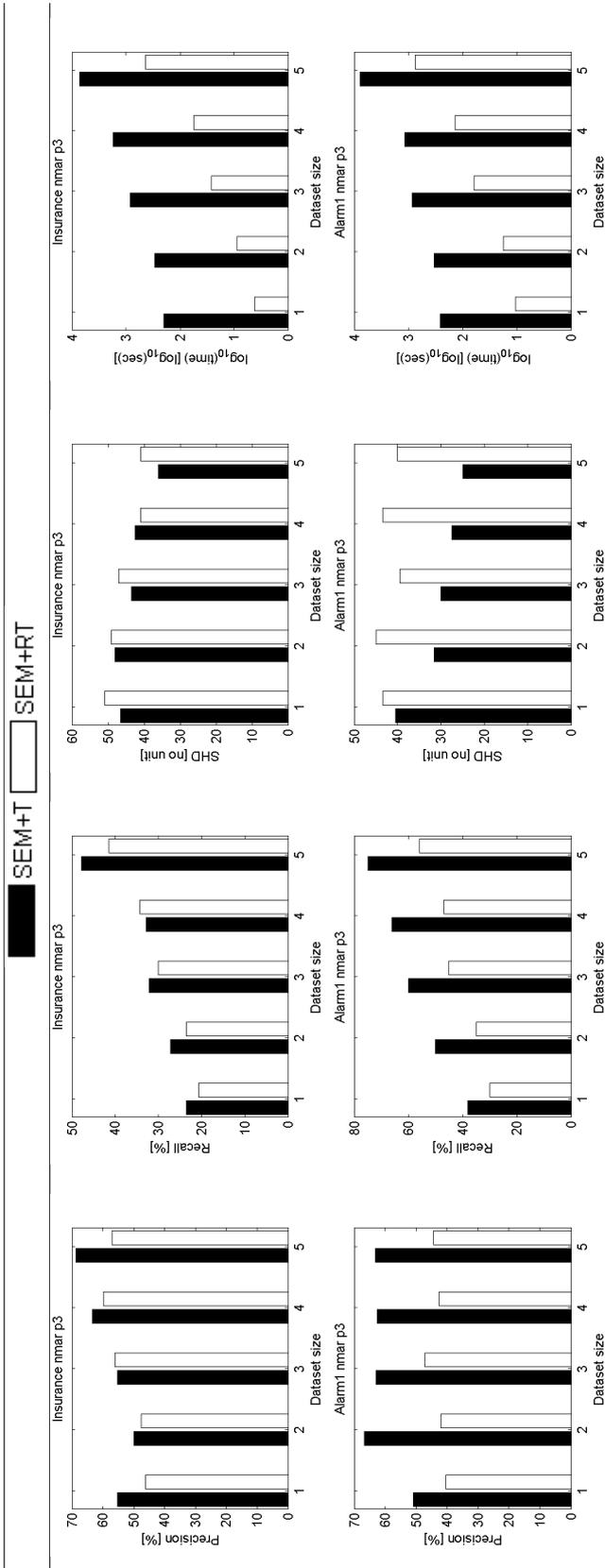


Precision-Recall-SHD-Tempo Asia e Child - nmar p3

SEM+T SEM+RT

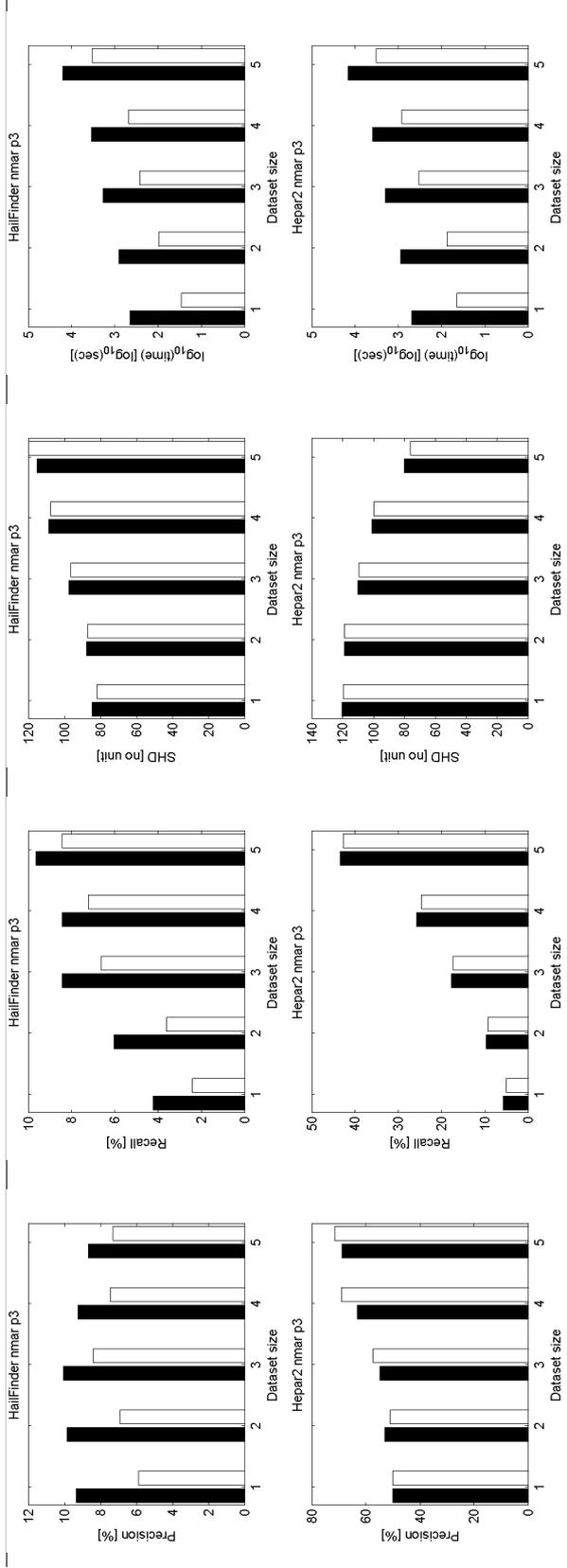


Precision-Recall-SHD-Tempo Insurance e Alarm1 - nmar p3



Precision-Recall-SHD-Tempo HaiFinder e Hepar2 - nmar p3

SEM+T SEM+RT



# Conclusione

In questa tesi si è vista una comparazione tra i metodi esistenti in letteratura per l'apprendimento della struttura di una rete bayesiana a partire da dati incompleti per cui esistesse un'implementazione di pubblico dominio. Il problema è molto importante nella realtà, in quanto il caso di incompletezza nei dati rappresenta una situazione estremamente frequente, almeno per quanto riguarda domini come quello biologico o medico. Pertanto, essere in grado di ricostruire correttamente il DAG di una rete bayesiana a partire da dati incompleti consentirebbe un'estensione notevole alle possibilità applicative di tali strutture, che sono contemporaneamente tanto potenti per eseguire inferenze di tipo probabilistico quanto utili nella loro immediatezza espressiva.

Si è partiti introducendo le principali nozioni teoriche (probabilità e teoria dei grafi) a cui sono seguite tutte le definizioni necessarie sulle reti bayesiane. Si è poi proseguito presentando gli algoritmi oggetto di confronto e il setup sperimentale utilizzato. Si sono utilizzati quattro algoritmi (MMHC (Tsamardinos, Brown e Aliferis 2006); SEM (Friedman 1997) (Friedman 1998); GES-EM (Borchani, Amor e Khalfallah 2007); MWST-EM (Leray e Francois 2005)) in tredici varianti in base all'inizializzazione da rete vuota o da albero MWST ottenuto con MWST-EM ed ai diversi score disponibili per SEM. Infine, si sono presentate le tabelle ed i grafici relativi ai risultati delle misure utilizzate nel confronto, cioè: precisione, recall, Structural Hamming Distance, tempo e area sotto la curva (AUC) delle curve Precision vs. Recall.

I due metodi che sono risultati globalmente migliori dal confronto sono stati SEM con score BDeu ed inizializzazione con MWST-EM ed imputazione del dataset incompleto con k-nn seguita da MMHC. SEM+T ha spesso una precisione maggiore, k-nn gode di un migliore recall, entrambi hanno buone performance in base alla structural Hamming distance, mentre in base al tempo la preferenza va verso MMHC con k-nn. Si tratta comunque di considerazioni di massima, in quanto non è emerso un vincitore unico né tra i due, né rispetto agli altri algoritmi (ad esempio, anche SEM da rete vuota ha spesso buone prestazioni). Si è verificato come ad incidere sulla bontà dei risultati siano principalmente la cardinalità del dataset, la probabilità di missingness ed il numero di nodi della rete. Nessuna nota importante sembra invece avere il meccanismo che opera nella generazione di incompletezza dei dati. Per le quattro reti più piccole, precisione e recall variano tra il 40-60% con cardinalità del dataset pari a 100 e salgono anche fino all'80-100% per dataset di taglia 5000. I risultati di Insurance sono leggermente peggiori, ma ciò è probabilmente dovuto all'alto numero di parametri della rete. Le SHD si mantengono a livelli accettabili, fino a scendere (e talvolta azzerarsi) per i dataset più grandi. Le due reti più grandi mostrano invece risultati decisamente peggiori, con precisione e recall che si aggirano sul

10% (anche se Hepar2 risponde bene all'aumentare dei dati) e SHD decisamente superiori al numero di archi della rete originaria. Vale probabilmente la pena di precisare che gli stessi andamenti si registrano nel caso di dati completi, per cui non è corretto ritenere la sola presenza di incompletezza la principale responsabile del degrado delle prestazioni. Se ne potrebbe invece dedurre che, di fronte a reti troppo grandi per il numero di dati a disposizione, sia consigliabile ricorrere a metodi diversi dalle reti bayesiane.

k-nn+MMHC e SEM si sono dimostrati i metodi che uniscono le migliori prestazioni alla migliore velocità, per quanto la variante SEM+T aumenti la qualità a scapito dell'efficienza computazionale. GES-EM si è dimostrato un metodo affidabile ma estremamente lento, ciò aggravato dal fatto che le sue prestazioni sono comunque inferiori rispetto agli altri due algoritmi.

Per valutare se un approccio ensemble learning potesse dare vantaggi rispetto all'uso di un solo algoritmo, abbiamo utilizzato una strategia greedy basata sulla misura della AUC delle curve Precision vs. Recall (per le quattro reti più piccole). In generale, non è emersa una combinazione vincente e spesso l'aggiunta di altri metodi tende a peggiorare la AUC. Tuttavia, in circa il 40% dei casi si è visto che la combinazione selezionata era SEM+T o k-nn o knn assieme a SEM+T. Una verifica esaustiva su tutte le reti ha mostrato come l'ensemble learning aumenti l'affidabilità degli archi individuati rispetto al metodo peggiore, anche se le AUC tendono ad essere minori rispetto al metodo migliore. Considerato che però non si ha un modo a priori per capire quale dei due sia di volta in volta il migliore e dato che si tratta in entrambi i casi di metodi veloci, può valere la pena di incrociare i risultati.

Infine, tenendo conto che Friedman consiglia di inizializzare SEM con una rete casuale, si è ritenuto corretto eseguire delle prove con "gruppo di controllo". Infatti, ci si può ragionevolmente attendere che l'ottimo risultato di SEM+T sia semplicemente dovuto alla partenza da rete non nulla piuttosto che ad un reale vantaggio dato da MWST-EM. SEM+T non ha prodotto risultati di qualità significativamente più alta rispetto all'inizializzazione di SEM con albero casuale, ad eccezione di Alarm ed in parte delle reti più grandi, nelle quali comunque la qualità è di per sé bassissima. Il vantaggio in termini di tempo, invece, favorisce notevolmente l'inizializzazione casuale.

Per quanto, al meglio delle nostre conoscenze, questo studio sia unico nel suo genere per dimensioni (4 algoritmi utilizzati in 13 varianti, 6 reti di prova da 8 a 70 nodi, 29700 reti apprese, circa 4 anni di tempo-macchina single-CPU) esso presenta notevoli limiti:

1. Mancano alcuni metodi. Per quanto non ci siano molti metodi stato-dell'arte per l'apprendimento di struttura in caso di dati incompleti, qualche algoritmo interessante esiste e qui non è stato trattato. In particolare, si sta provvedendo ad una implementazione di BSEM (Elidan 2011), una variante bagged di SEM particolarmente recente e apparentemente promettente, a partire dall'implementazione di SEM del pacchetto BNT-SLP.
2. Le reti di prova utilizzate sono molto sparse. Evidentemente, questo non rispecchia necessariamente un caso reale e meriterebbe un approfondimento. Infatti, è possibile che l'inizializzazione di SEM con un albero dia ottimi risultati in quanto la struttura di partenza è già blandamente simile alla rete reale, cosa che non è detto accada con reti dense.

3. Gli algoritmi sono implementati con linguaggi diversi, alterando una realistica possibilità di confronto temporale (e potenzialmente anche di performance, qualora, ad esempio, aumentare le iterazioni di GES-EM portasse a risultati migliori rispetto ai concorrenti in un lasso di tempo simile).
4. I meccanismi di generazione della missingness utilizzati sono semplicistici rispetto a quanto può accadere nella realtà (in cui, ad esempio, più variabili potrebbero influenzarne un'altra ed in base a ciò i valori che essa assume potrebbero condizionare la mancanza del proprio dato, secondo uno schema ben più complesso di quanto da noi utilizzato). Potrebbe essere interessante eseguire ulteriori approfondimenti con schemi di incompletezza più elaborati.

Nel futuro, sarebbe interessante provvedere ad una comparazione dei metodi qui studiati, ed eventualmente anche di altri, eseguita a partire da dati reali di varia natura, diversi domini e differente numero di variabili e dati. In questo modo, si potrebbero incrociare i risultati ottenuti da dataset reali con le valutazioni emerse a partire dallo studio teorico.



# Ringraziamenti

Alla Dott.ssa Di Camillo ed al Dott. Sambo per la rapidità con cui si sono prodigati per rispondere ai miei molti dubbi e per la disponibilità manifestatami, in senso accademico, ma non solo.

Ai miei genitori, che hanno sempre cercato di essermi di esempio e mi hanno sostenuto durante questo lungo percorso universitario, nonostante la difficoltà di capire il mio strano modo di vedere il mondo.

Grazie, Arianna, sorella testardamente occupata nell'essere di sostegno al suo strambo fratello. Sinceramente, in questo ringraziamento c'è più di quanto possa essere ufficialmente espresso.



# Bibliografia

- Akaike, H. (1973). «Information theory and an extension of the maximum likelihood principle». In: *Second International Symposium on Information Theory*. A cura di B. N. Petrov e B. F. Csaki. Akademiai Kiado. Budapest, pp. 267–281.
- Andreassen, S. et al. (1991). «A Model-based Approach to Insulin Adjustment». In: *Proceedings of the 3rd Conference of Artificial Intelligence in Medicine*. Springer-Verlag, pp. 239–248.
- Badaloni, S., F. Sambo e F. Venco (In press). «Bayesian Network structure learning: hybridizing complete search with independence tests». In: *AI Communications*.
- Baldi, P. (1998). *Calcolo delle probabilità e statistica*. Milano: McGraw-Hill.
- Beinlich, I. A. et al. (1989). «The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks». In: *Second European Conference in Artificial Intelligence in Medicine*.
- Bertsekas, D. P. e J. N. Tsitsikilis (2008). *Introduction to probability, 2nd edition*. Nashua, NH: Athena Scientific.
- Binder, J. et al. (1997). «Adaptive Probabilistic Networks with Hidden Variables». In: *Machine Learning* 29.2–3, pp. 213–244. URL: <http://www.cs.huji.ac.il/site/labs/compbio/Repository/Datasets/insurance/insurance.htm>.
- Borchani, H., N. B. Amor e F. Khalfallah (2007). «Learning and Evaluating Bayesian Network Equivalence Classes from Incomplete Data». In: *International Journal of Pattern Recognition and Artificial Intelligence*.
- Breiman, L. (1996). «Bagging predictors». In: *Machine Learning*.
- Cheeseman, P. e J. Stutz (1995). «Bayesian Classification (Autoclass): Theory and Results». In: *Advances in Knowledge Discovery and Data Mining*. A cura di D. Fayyad et al. Menlo Park, California: AAAI Press.
- Cheng, J., D. Bell e W. Liu (1997). «Learning belief networks from data: An information theory based approach». In: *Proceedings of the sixth ACM International Conference on Information and Knowledge Management*.
- Cheng, J. et al. (2002a). «Learning Bayesian networks from data: An information-theory based approach». In: *Artificial Intelligence* 137, pp. 43–90.
- Cheng, J. et al. (2002b). «Learning bayesian networks from data: an information-theory based approach». In: *The Artificial Intelligence Journal* 137, pp. 43–90.
- Chickering, D. M. (2002). «Optimal Structure Identification With Greedy Search». In: *Journal of Machine Learning Research* 3, pp. 507–554.

- Chow, C. e C. Liu (1968). «Approximating discrete probability distributions with dependency trees». In: *IEEE Transactions on Information Theory* 14, pp. 462–467.
- Cooper, G. F. e E. Herskovits (1992). «A Bayesian Method for the Induction of Probabilistic Networks from Data». In: *Machine Learning* 9.
- Cormen, T. H. et al. (2005). *Introduzione agli algoritmi e strutture dati, 2a edizione*. trad. it. Colussi L. Milano: McGraw-Hill.
- Cowell, R. G. et al. (1999). *Probabilistic networks and expert systems*. Springer.
- Dempster, A., N. Laird e D. Rubin (1977). «Maximum likelihood from incomplete data via the EM algorithm». In: *Journal of the Royal Statistical Society* 39, 1:38.
- Dor, D. e M. Tarsi (1992). *A simple algorithm to construct a consistent extension of a partially oriented graph*. R-185. Cognitive Systems Laboratory, UCLA Computer Science Department.
- Elidan, G. (2011). «Bagged Structure Learning of Bayesian Networks». In: *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AI-Stats)*, pp. 251–259.
- Friedman, N. (1997). «Learning Belief Networks in the Presence of Missing Values and Hidden Variables». In: *Proceedings of the Fourteenth International Conference on Machine Learning*. ICML '97. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., pp. 125–133. ISBN: 1-55860-486-3. URL: <http://dl.acm.org/citation.cfm?id=645526.657145>.
- (1998). «The Bayesian Structural EM Algorithm». In: *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*. UAI'98. Madison, Wisconsin: Morgan Kaufmann Publishers Inc., pp. 129–138. ISBN: 1-55860-555-X. URL: <http://dl.acm.org/citation.cfm?id=2074094.2074110>.
- Friedman, N., I. Nachman e D. Pe'er (1999). «Learning Bayesian network structure from massive datasets: "The sparse candidate" algorithm». In: *Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI-99)*.
- Garcia-Laencina, P. J. e J.-L. Sancho-Gómez (2010). «Pattern classification with missing data: a review». In: *Neural Computation & Application* 19, pp. 263–282.
- Geman, S. e D. Geman (1984). «Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images». In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, pp. 721–742.
- Goodrich, M. T. e R. Tamassia (2006). *Data structures and Algorithms in Java, 4th edition*. Hoboken, NJ: John Wiley & Sons.
- Heckerman, D., D. Geiger e D. M. Chickering (1995). «Learning Bayesian Networks: The Combination of Knowledge and Statistical Data». In: *Machine Learning* 20, pp. 197–243.
- Jensen, A. e F. Jensen (1996). «Midas - An influence diagram for management of mildew in winter wheat». In: *Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence (UAI-96)*. Morgan Kaufmann Publishers, pp. 349–356.
- Kocka, T. e R. Castelo (2001). «Improved learning of Bayesian Networks». In: *Proceedings of the Conference on Uncertainty in artificial intelligence*. A cura di D. Koller e J. Breese, pp. 269–276.

- Kristensen, K. e I. A. Rasmussen (2002). «The use of a Bayesian Network in the design of a decision support system for growing malting barley without use of pesticides». In: *Computers and Electronics in Agriculture* 33, pp. 197–217.
- Kruskal, J. B. Jr. (1956). «On the shortest spanning subtree of a graph and the traveling salesman problem». In: *Procedures of the American Mathematical Society* 7, pp. 48–50.
- Lauritzen, S. L. e D. J. Spiegelhalter (1988). «Local Computation with Probabilities on Graphical Structures and their Application to Expert Systems (with discussion)». In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 50.2, pp. 157–224.
- Leray, P. e O. Francois (2004). *BNT structure learning package: Documentation and experiments*.
- (2005). *Bayesian Network Structural Learning and Incomplete Data*. Espo, Finland.
- Lin, J.-H. e P. J. Haug (2008). «Exploiting missing clinical data in Bayesian network modeling for predicting medical problems». In: *Journal of Biomedical Informatics*, pp. 1–14.
- Little, Roderick J. A. e Donald B. Rubin (1986). *Statistical Analysis with Missing Data*. New York, NY, USA: John Wiley & Sons, Inc. ISBN: 0-471-80254-9.
- Margaritis, D. (2003). «Learning Bayesian Network Model Structure from Data». Available as Technical Report CMU-CS-03-153. Tesi di dott. Pittsburgh, PA: School of Computer Science, Carnegie-Mellon University.
- Meek, C. (1997). «Graphical Models: Selecting causal and statistical models». Tesi di dott. Carnegie Mellon University.
- Moore, A. e W. Wong (2003). «Optimal reinsertion: a new search operator for accelerated and more accurate Bayesian network structure learning». In: *Twentieth International Conference on Machine Learning (ICML-2003)*.
- Murphy, K. «The BayesNet Toolbox for Matlab». In: *Computing Science and Statistics: Proceedings of Interface*. Vol. 33.
- Neapolitan, R. E. (2003). *Learning Bayesian Networks*. Prentice-Hall.
- Onisko, A. (2003). «Probabilistic Causal Models in Medicine: Application to Diagnosis of Liver Disorders.» Tesi di dott. Warsaw: Institute of Biocybernetics e Biomedical Engineering, Polish Academy of Science.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.
- Pearl, J. e T. S. Verma (1991). «A theory of inferred causation». In: *Principles of Knowledge Representation and Reasoning*. A cura di J. F. Allen, R. Fikes e E. Sandewall. Morgan Kaufmann Publishers, 441–452.
- Ramoni, M. e P. Sebastiani (1997). «Learning Bayesian Networks from Incomplete Databases». In: *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., pp. 401–408.
- Riggelsen, C. (2006). «Learning Bayesian Networks from incomplete Data: An Efficient Method for Generating Approximate Predictive Distributions». In: SIAM.
- Riggelsen, Carsten e Ad Feelders (2005). «Learning Bayesian network models from incomplete data using importance sampling». In: *Proc. of Artificial Intelligence and Statistics*, pp. 301–308.
- Rissanen, J. (1987). «Stochastic complexity (with discussion)». In: *Journal of the Royal Statistical Society* 49, 223–239 and 253–265.

- Russell, S. e P. Norvig (2003). *Artificial Intelligence: A Modern Approach, 2nd edition*. Pearson Education.
- Schwarz, G. (1978). «Estimating the dimension of a model». In: *The Annals of Statistics* 2.6, pp. 461–464.
- Scutari, M. (2009). *Learning Bayesian Networks with the bnlearn R package*.
- Spiegelhalter, D. e S. Lauritzen (1990). «Sequential updating of conditional probabilities on directed graphical structures». In: *Networks* 20, pp. 579–605.
- Spirites, P., G. Glymour e R. Scheines (2000). *Causation, Prediction, and Search*. New York: Springer-Verlag.
- Tanner, M. A. e W. H. Wong (1987). «The calculation of posterior distributions by data augmentation». In: vol. 82. 398. Taylor & Francis, pp. 528–540.
- Tean, R Core et al. (2012). «R: A language and environment for statistical computing». In:
- Tian, F., H. Zang e Y. Lu (2003). «Learning Bayesian Networks from Incomplete Data Based on EMI Method». In: *Proceedings of the third IEEE International Conference on Data Mining*.
- Tsamardinos, I., C. F. Aliferis e A. Statnikov (2003). «Time and Sample Efficient Discovery of Markov Blankets and Direct Causal Relations». In: *KDD '03: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, pp. 673–678.
- Tsamardinos, I., L. E. Brown e C. F. Aliferis (2006). «The max-min hill-climbing Bayesian network structure learning algorithm». In: *Machine Learning* 65, pp. 31–78.
- Wong, M. L. e Y. Y. Guo (2008). «Learning Bayesian Networks from incomplete databases using a novel evolutionary algorithm». In: *Decision Support Systems* 45, pp. 368–383.
- Wong, M. L. e K. S. Leung (2004). «An Efficient Data Mining Method for Learning Bayesian Networks Using an Evolutionary Algorithm-Based Hybrid Approach». In: *IEEE Transactions on Evolutionary Computation* 8.4, pp. 378–404.

# Indice analitico

- Albero libero, 15
- Albero radicato, 15
- Apprendimento di parametri, 27
- Apprendimento di strutture, 33
- Approccio constraint based, 33
- Approccio score based, 33
- Available Cases Analysis, 37
  
- B-spazio, 45
- Bayes, Legge di, 6, 12
- Bayesian Information Criterion (BIC), 34
  
- Cammini attivi e bloccati, 19
- Campione binomiale, 29
- Catena (nodi), 15
- Causa, 17
- Coda (nodi), 15
- Complete Cases Analysis, 37
- CPDAG, 20
  
- d-separazione, 19
- d-separazione per classi di equivalenza, 20
- d-separazione, Identificazione di, 19
- DAG, 15
- DAG causale, 17
- DAG, Schema della classe di equivalenza, 20
- Densità condizionata, 11
- Deviazione standard, 9
- Dipendenza diretta, 21
- Distribuzione beta, 29
  
- E-spazio, 45
- EM, 32
- Equivalent Sample Size, 32
- Equivalenza di indipendenze condizionali, 19
- ESS, 32
- Evento, 4
- Expectation-Maximization, 32
  
- Faithfulness, 21
- Famiglia coniugata di funzioni di densità, 30
- Funzione beta, 29
- Funzione di densità di Dirichlet, 30
- Funzione di distribuzione discreta, 8
- Funzione di ripartizione, 9
- Funzione gamma, 28
  
- Grafo, 13
- Grafo aciclico, 14
- Grafo Diretto Aciclico, 15
- Grafo, Archi di, 13
- Grafo, Cammino di un, 14
- Grafo, componenti connesse di un, 14
- Grafo, Vertici di, 13
- Greedy Equivalence Search-Expectation Maximization (GES-EM), 44
  
- Heterogeneous Euclidean Overlap Metric (HEOM), 38
  
- i.i.d., 28
- Indifferenza, Principio di, 5
- Indipendenza condizionata, 12
- Indipendenza condizionata di eventi, 7
- Indipendenza di eventi, 7
- Indipendenza di un insieme di eventi, 7
- Indipendenza di v.a., 12
- Indipendenza di v.a. da un evento, 12
- Indipendenza indotta, 19
- Inferenza, 22
- Iperparametri, 29, 30
  
- k-nn, 37
- Kolmogorov, assiomi di, 4
  
- Link (nodi), 15
  
- Manipolazione, 17
- MAP, 28
- Mappa perfetta, 21

- Markov blanket, 21
- Markov boundary, 22
- Markov, Assunzione di causalità, 18
- Markov, Condizione di, 16
- Markov, Equivalenza di, 20
- Massima Verosimiglianza, 28
- Matrice indicatrice dei dati mancanti, 26
- Maximum A Posteriori, 28
- Maximum Likelihood, 28
- Maximum Weight Spanning Tree, 43
- Maximum Weight Spanning Tree-Expectation Maximization (MWST-EM), 43
- Minimalità, 21
- Missing At Random, 26
- Missing Completely At Random, 26
- Missingness, 25
- Missingness ignorabile, 26
- ML, 28
- Modello probabilistico, 4
- Momento di una v.a., 9
- Not Missing At Random, 26
- Osservazione, 13
- PDAG, 20
- Posterior bayesiana, 13
- Prior bayesiano, 13
- Probabilità, 4
- Probabilità condizionata, 6
- Probabilità totale, Teorema della, 6
- Probabilità, approcci alla, 5
- Probabilità, Densità congiunta di, 10
- Probabilità, Densità di, 8
- Probabilità, Densità marginale di, 10
- Probabilità, legge di, 4
- Probabilità, Spazio di, 4
- Probabilità, Teoria delle, 3
- Regola del prodotto, 6
- Regola della catena per v.a., 11
- Rete bayesiana, 17
- Rete bayesiana, Parametri di, 17
- Rete bayesiana, Struttura di, 17
- Scheletro di un DAG, 20
- Score, 33
- Spazio campionario, 4
- Spazio degli eventi, 4
- Structural Expectation Maximization (SEM), 42
- Structural Hamming Distance (SHD), 57
- Testa (nodi), 15
- Training set, 25
- Valore atteso, 9
- Variabile aleatoria, 8
- Variabile aleatoria n-dimensionale, 10
- Variabile aleatoria, Spazio di una, 8
- Varianza, 9