

## L'algoritmo P-SBAS per l'elaborazione automatica di dati SAR interferometrici su grande scala

Claudio De Luca (\*,\*\*), Stefano Elefante (\*), Pasquale Imperatore (\*), Michele Manunta (\*), Ivana Zinno (\*), Valentina Casola (\*\*), Francesco Casu (\*)

(\*) CNR IREA, Via Diocleziano 328, 80124 Napoli, Italy

(\*\*) Università degli Studi di Napoli "Federico II", DIETI, Via Claudio 21, 80124, Napoli

### Riassunto

L'algoritmo di telerilevamento SBAS permette di generare mappe e serie storiche di deformazione del suolo con accuratze centimetriche e rappresenta ormai un valido strumento per lo studio dei complessi fenomeni geofisici che caratterizzano vaste aree della superficie terrestre. In questo lavoro è presentata una soluzione parallela dell'algoritmo SBAS, chiamata P-SBAS, orientata all'elaborazione in tempi contenuti di grandi archivi di dati SAR, sfruttando il parallelismo offerto dalle moderne infrastrutture di calcolo. Al fine di dimostrare l'efficacia della soluzione algoritmica proposta, è stata inoltre realizzata un'analisi delle prestazioni dell'algoritmo P-SBAS tramite una valutazione di alcune metriche rappresentative (*speedup*, bilanciamento del carico, efficienza, ...). L'analisi delle prestazioni, condotta su due data set reali di immagini SAR (acquisite dai sensori ENVISAT e COSMO-SkyMed) e sfruttando una piattaforma di calcolo con 32 processori, ha infine mostrato l'efficacia della soluzione P-SBAS proposta.

### Abstract

The remote sensing SBAS (Small BAseline Subset) algorithm allows to evaluate mean velocity and time series deformation maps of the ground with centimeter accuracy, which represent a valuable tool for the analysis of the complex phenomena characterizing the surface deformation dynamics of Earth large areas. In this paper, we show a parallel computing solution for the SBAS processing chain, named P-SBAS. Such a solution is oriented to the effective exploitation of the parallelism offered by modern computational platform to process the available huge SAR data archives in reasonable time-frames. The quantitative evaluation of the computational efficiency of the implemented parallel solution in terms of achieved speedup is addressed to demonstrate the effectiveness of the proposed approach. An experimental analysis has been carried out on real data by employing a computational platform comprising 32 processors. In particular, a performance analysis has been conducted by exploiting different SAR datasets pertinent to different sensors (ENVISAT and COSMO-Sky-Med).

### 1. Introduzione

L'interferometria SAR differenziale (DInSAR) è una tecnica di telerilevamento per la stima delle deformazioni del suolo con accuratze centimetriche e in alcuni casi millimetriche. (Goldstein, Zebker, 1989) Negli ultimi decenni, grazie ad una sempre crescente disponibilità di dati SAR satellitari e allo sviluppo di algoritmi DInSAR avanzati come la tecnica SBAS (Small Baseline Subset) (Berardino et al., 2002), l'interferometria differenziale si è affermata all'interno della comunità geofisica (vulcanologia, sismologia, etc.) come un efficace strumento di telerilevamento non solo per comprendere meglio i fenomeni geofisici a scala locale e regionale, ma anche per dare supporto alla gestione e mitigazione dei rischi naturali e antropici.

Lo scenario attuale del DInSAR è caratterizzato dalla disponibilità di grossi archivi di dati SAR satellitari acquisiti fin dai primi anni '90 dai sensori ERS, ENVISAT e RADARSAT operanti in banda C e, successivamente, dalle costellazioni che operano in banda X come COSMO-SkyMed (CSK) e TerraSAR-X (TSX). Nel prossimo futuro inoltre i sensori SAR a bordo della costellazione europea Sentinel-1 permetteranno di compiere un ulteriore passo in avanti nella disponibilità di dati acquisiti e nella capacità di monitoraggio ed analisi delle deformazioni del suolo, grazie ad una politica di acquisizione *global coverage* e un tempo di rivisita di 12 giorni che scenderà a 6 giorni con la contemporanea presenza dei sensori Sentinel-1A e 1B (Snoeij et al., 2010).

In un contesto DInSAR, quindi, al fine di poter beneficiare di questa enorme mole di dati limitando i tempi necessari per le elaborazioni interferometriche, è opportuno disporre non solo di risorse di calcolo ad alte prestazioni (come piattaforme *cluster*, *grid* o *cloud*), ma anche di algoritmi DInSAR che siano in grado di sfruttarle in modo efficace ed efficiente. A tal proposito, in questo lavoro si presenta una versione parallela dell'algoritmo SBAS (P-SBAS) che sfrutta le potenzialità offerte dalle moderne architetture di calcolo parallelo e che in generale si basa su un approccio a due livelli di parallelizzazione (Elefante et al., 2013) (Casu et al., 2014).

Nel seguito, dopo la descrizione dell'algoritmo P-SBAS, è presentata un'analisi sperimentale delle prestazioni della soluzione proposta, condotta su dati SAR reali acquisiti dai sensori ENVISAT e COSMO-SkyMed. In particolare, l'analisi delle prestazioni si basa sull'utilizzo di alcune metriche che consentono di valutare il comportamento dell'algoritmo rispetto al numero di processori (CPUs) utilizzati (scalabilità), al variare dello scenario applicativo.

## 2. Descrizione dell'algoritmo P-SBAS

In questo paragrafo è data una breve descrizione dell'algoritmo P-SBAS proposto; per maggiori dettagli si rimanda a (Casu et al., 2014). Data la sua complessità, in Figura 1 è mostrato il diagramma a blocchi ad alto livello dell'algoritmo P-SBAS. In particolare, i blocchi che sono eseguiti in modo parallelo sono raffigurati in rosso/multi-livello, mentre quelli neri rappresentano i passi sequenziali della catena di elaborazione.

I blocchi A e B realizzano rispettivamente la focalizzazione dei dati grezzi e la conversione del *Digital Elevation Model* (DEM) sulla geometria di acquisizione del sensore SAR. I blocchi da C a G realizzano rispettivamente, la registrazione geometrica delle immagini (Sansosti et al., 2006), la selezione delle coppie interferometriche, il raffinamento della registrazione effettuata al passo C, la stima dei parametri interferometrici e la generazione degli interferogrammi differenziali. La fase di ciascun interferogramma è disponibile nell'intervallo ambiguo  $[-\pi, \pi]$  (fase arrotondata), mentre per poter essere utilizzata nel calcolo della serie storica di deformazione deve essere nota nella sua completa evoluzione (fase srotolata). A tal fine, i passi H e I risolvono l'ambiguità della fase arrotondata implementando una versione parallela dell'algoritmo di *Phase Unwrapping* (srotolamento della fase) noto come EMCF (Pepe, Lanari, 2006). Una inversione del sistema di equazioni delle fase srotolate (blocco J) e alcune successive operazioni di filtraggio (blocco L) permettono di generare la serie storica di deformazione. Opzionalmente, è possibile eseguire una procedura di rimozione degli errori orbitali residui (blocco K) che implica un'ulteriore esecuzione delle operazioni di *Phase Unwrapping* (H, I) e di generazione della serie di deformazione (J, L).

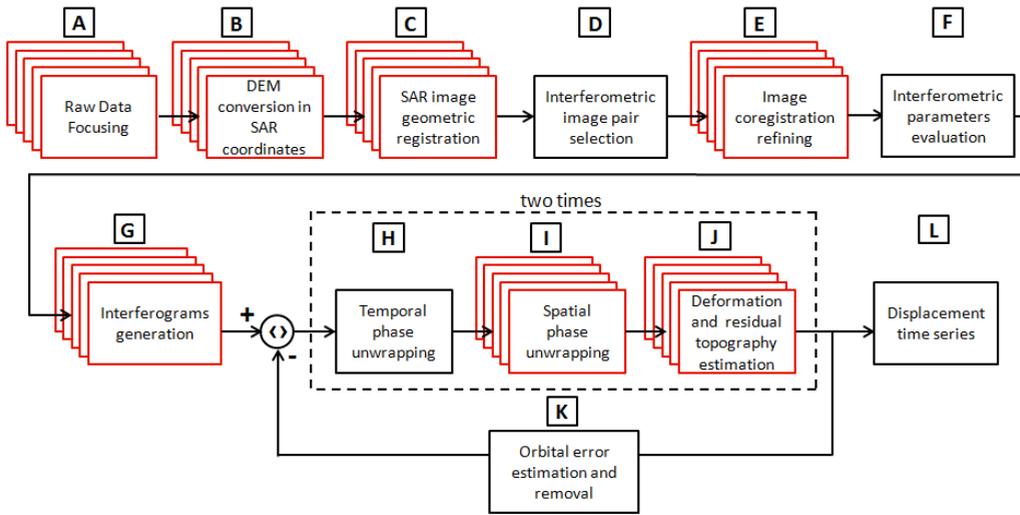


Figura 1 - Diagramma a blocchi dell'algoritmo P-SBAS. In rosso i blocchi implementati con logica parallela.

Vale la pena enfatizzare che per la realizzazione di tale algoritmo sono state adottate differenti strategie di parallelizzazione. In particolare, sono state sfruttate le possibilità di *processing* multi-nodo e multi-*thread* delle moderne architetture di calcolo parallelo e sono quindi state implementate delle procedure che le riescono a sfruttare al meglio, in alcuni casi anche in maniera combinata, in dipendenza dal passo di elaborazione dell'algoritmo SBAS (Casu et al., 2014).

### 3. Descrizione dei datasets

La baia di Napoli è stata selezionata come area di *test* per l'analisi delle prestazioni dell'algoritmo P-SBAS, non solo perché rappresenta un interessante *test site* interferometrico per la presenza della caldera dei Campi Flegrei, del Vesuvio e della città di Napoli, ma anche perché erano disponibili due data set di immagini SAR acquisite da sistemi satellitari differenti.

Il primo data set è formato da 64 acquisizioni ASAR-ENVISAT, acquisite su orbita ascendente tra il novembre 2002 e settembre 2010. In particolare in questo caso sono stati generati 195 interferogrammi differenziali caratterizzati da valori di baseline (la distanza fra le orbite delle acquisizioni della coppia interferometrica) perpendicolare più piccoli di 400 m e da valori di baseline temporale di circa 1500 giorni. Il secondo data set è a sua volta formato da 64 immagini COSMO-SkyMed acquisite su orbita ascendente nel periodo che va dal gennaio 2012 al febbraio 2014. Anche per quest'ultimo data set sono stati generati 195 interferogrammi differenziali, caratterizzati da una baseline perpendicolare minore di 800 m e da una baseline temporale massima di 400 giorni. Le principali caratteristiche dei data set di immagini SAR utilizzati sono raccolte in Tabella. 1.

Tabella 1 - Parametri Caratteristici dei data set SAR utilizzati.

	ENVISAT	COSMO-SkyMed
Dimensione dell'area analizzata	100x100 km <sup>2</sup>	60x60 km <sup>2</sup>
Numero di acquisizioni SAR	64	64
Dimensione Immagini [pixel]	18000x5000	19000x20000
Numero di interferogrammi	195	195
Intervallo Temporale	nov. 2002 – set. 2010	gen. 2012 – feb. 2014

Tabella 2 - Caratteristiche della Piattaforma di Calcolo.

	Caratteristiche
Sistema Operativo	GNU/Linux 2.6.32
Processore	8-core CPU – 2.6 GHz Intel Xeon E5-2670
Numero di nodi	16 (2 processori per nodo)
Memoria	384 GB di RAM
Interconnessione	40 Gb/s InfiniBand

#### 4. Analisi delle prestazioni e risultati interferometrici

È stata condotta un'analisi delle prestazioni dell'algoritmo P-SBAS in termini di *speedup*, al variare del numero di processori fissata la dimensione del problema. Lo *speedup*  $S_N$  di un programma parallelo e con tempo di esecuzione  $T_N$  è definito come:

$$S_N = \frac{T_1}{T_N} \quad [1]$$

Dove  $N$  è il numero di processori utilizzati per il problema considerato,  $T_N$  è il tempo di esecuzione del programma parallelo utilizzando  $N$  processori e  $T_1$  è il tempo di esecuzione del medesimo programma sfruttando un singolo processore (equivalente ad una configurazione sequenziale) (Hager, Wellein, 2010) (El-Rewini, Abd-El-Barr, 2005).

Vale la pena sottolineare che il carico di lavoro (*workload*) relativo ai differenti passi della catena P-SBAS dipende significativamente dalla dimensione del data set utilizzato per l'elaborazione (in termini di numero di acquisizioni, area coperta, ...). Infatti, i parametri di ingresso che condizionano il carico computazionale sono rappresentati dal: numero di immagini del data set, numero di pixel per ogni immagine, numero di interferogrammi generati e numero di pixel sui quali effettuare le operazioni di *Phase Unwrapping* (punti coerenti). Tali parametri impattano in maniera differente in relazione al passo algoritmico da effettuare. Ad esempio, i passi della catena che vanno dalla focalizzazione dei dati grezzi (blocco A in Figura 1) fino al blocco di generazione degli interferogrammi (blocco G in Figura 1) hanno un carico computazionale fortemente dipendente dal numero di immagini del data set, dal numero di interferogrammi generati e dalla loro dimensione (in pixel). Mentre la complessità dei passi che vanno dal *phase unwrapping* temporale (blocco H in Figura 1) fino alla generazione della serie storica di deformazione (blocco L in Figura 1), dipendono maggiormente dal numero di interferogrammi generati e dal numero di punti coerenti da srotolare. Si noti che il numero di punti coerenti è dipendente dalle caratteristiche elettromagnetiche della scena osservata in relazione al sensore utilizzato.

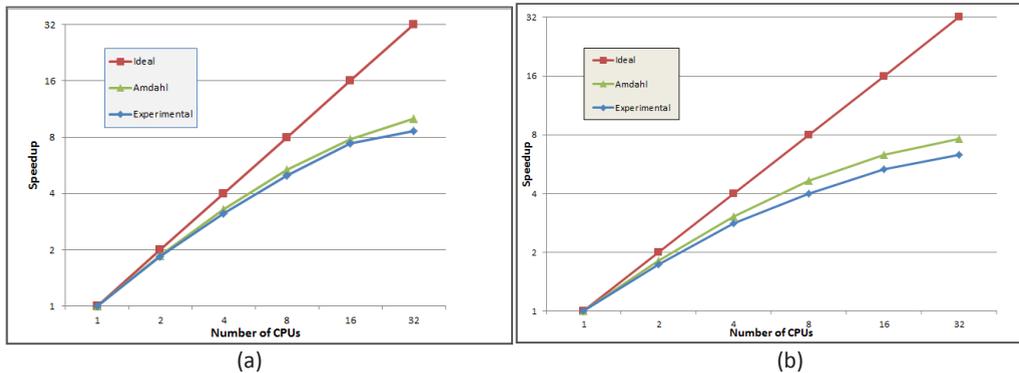


Figura 2 - Analisi delle prestazioni dell' algoritmo P-SBAS. (a) speedup in funzione del numero di processori utilizzati relativo al data set ENVISAT. Sono mostrati i risultati sperimentali (blu/rombi), il comportamento lineare ideale (rosso/quadrati), l'andamento secondo la legge di Amdahl (verde/triangoli). (b) come in (a) relativo al dataset COSMO-SkyMed.

I test sono stati condotti utilizzando una piattaforma di calcolo (*cluster*) costituita da 16 nodi, ognuno dei quali equipaggiato con 2 CPU (8-core 2.6 GHz Intel Xeon E5-2670) e 384 GB di RAM. Il *cluster* in questione è configurato con *storage* condiviso attraverso un Network File System (NFS) e utilizza una interconnessione tra i nodi con tecnologia InfiniBand da 40 Gb/s (Tabella 2). In modo da apprezzare quantitativamente i benefici della parallelizzazione realizzata con la soluzione P-SBAS è stata adottata una metrica appropriata.

In Figura 2 è mostrata l'analisi di scalabilità dell'algoritmo per entrambi i casi di studio (dati ENVISAT e COSMO-SkyMed). Tale analisi è stata condotta fissando la dimensione del problema (in termini di immagini e area studiata) al variare del numero di processori (CPUs) utilizzati dall'algoritmo. Per entrambi i casi di studio lo *speedup* è valutato in funzione del numero di processori  $N$  utilizzati da P-SBAS, ed in particolare sono messi in evidenza, i risultati sperimentali (blue/rombi), l'andamento ideale (rosso/quadrati), e l'andamento secondo la legge di Amdahl (verde/triangoli).

I risultati ottenuti mostrano valori di *speedup* vicini al limite di Amdahl, dovuti anche ad un buon bilanciamento del carico di lavoro dei processori. In particolare, il tempo di elaborazione del data set ENVISAT utilizzando P-SBAS su un unico processore è stato di circa 53 ore (2.2 giorni), rispetto alle 6 ore ottenute utilizzando 32 processori, lo *speedup* corrispondente è di circa 8.6. Per l'elaborazione del dataset COSMO-SkyMed, invece, utilizzando sempre 32 processori, lo *speedup* ottenuto è di 6.3, come mostrato in Figura 2.

La percentuale di tempo di elaborazione relativa alla sola parte sequenziale dell'algoritmo è stata stimata al 5% per il caso ENVISAT e al 10% per il caso COSMO-SkyMed. È giusto sottolineare che le differenze rilevate tra lo *speedup* dell'analisi sperimentale e quello valutato secondo la legge di Amdahl può essere giustificato come un residuo di carico di lavoro non bilanciato (Zinno et al., 2013) (Hager, Wellein, 2010). Inoltre i risultati ottenuti, coerentemente alla legge di Amdahl, mostrano che la maggiore sorgente di inefficienza nella soluzione algoritmica adottata è attribuibile alla presenza di una porzione sequenziale non trascurabile dell'algoritmo attuale. Tale porzione di codice, anche aumentando il numero di processori, vincola inferiormente il tempo totale di esecuzione e superiormente lo *speedup* massimo raggiungibile.

Infine, in Figura 3(a) e 3(b) sono mostrate le mappe di velocità media di deformazione sulla baia di Napoli ottenute rispettivamente dai dati ENVISAT e COSMO-Sky-Med elaborati nei test precedentemente presentati. Tali risultati sono in accordo con precedenti studi interferometrici sulla stessa area (Bonano et al., 2012) (Bonano et al., 2013) dimostrando anche la validità interferometrica dell'algoritmo P-SBAS proposto.

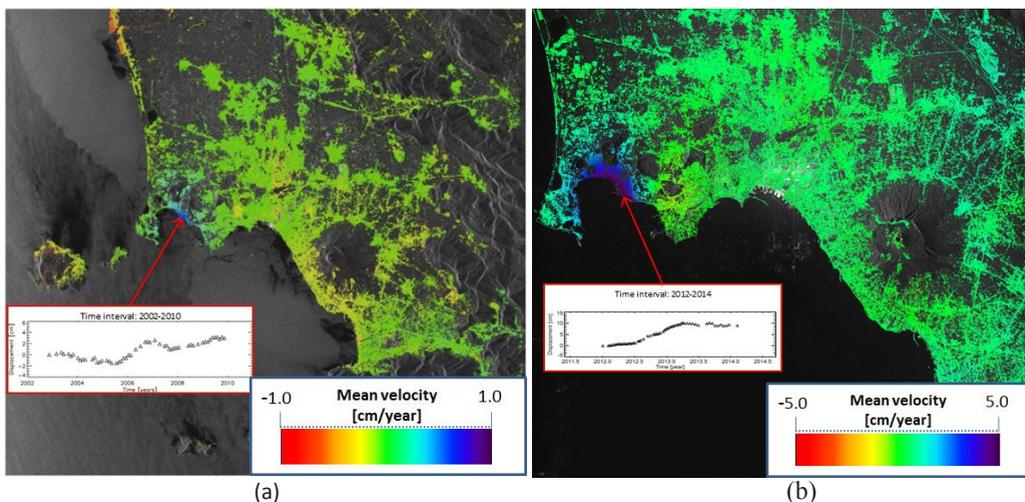


Figura 3 - Mappa di velocità media di deformazione prodotta con l'algoritmo P-SBAS e relativa alla baia di Napoli, ottenuta con dati ENVISAT (a) e COSMO-SkyMed (b). Per ciascuna mappa è anche mostrato il grafico della serie storica di deformazione relativa ad un pixel posizionato nell'area di massima deformazione.

## 5. Conclusioni

In questo lavoro è stata presentata una soluzione parallela dell'algoritmo SBAS e ne è stata effettuata un'analisi di scalabilità per verificarne le prestazioni e la validità. Sono stati considerati due casi di studio relativi ad un data set di immagini ENVISAT ed uno di immagini COSMO-SkyMed. I risultati hanno mostrato che la soluzione algoritmica proposta può essere utilizzata per lo sfruttamento efficace di piattaforme di calcolo parallelo, e per una configurazione efficiente dell'architettura di calcolo da impiegare per elaborare un assegnato data set, in termini di risorse utilizzate e tempo di *processing*.

I promettenti risultati mostrano che la soluzione P-SBAS proposta permette di abbattere notevolmente i tempi necessari ad un'elaborazione DInSAR, con importanti ricadute nell'analisi degli spostamenti della superficie terrestre e nel monitoraggio dei fenomeni deformativi naturali ed antropici. Infine, l'approccio P-SBAS proposto può essere di notevole impatto anche nell'elaborazione DInSAR della grossa mole di dati SAR acquisiti nel prossimo futuro dai sensori Sentinel-1 del programma Copernicus.

## 6. Bibliografia

- Goldstein R, and Zebker H.A., "Mappings small elevation changes over large areas: Differential radar interferometry," *J. Geophys. Res.*, vol.94, no.B7, pp.9183-9191, 1989
- Berardino P., Fornaro G., Lanari R., and Sansosti E., "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2375-2383, Nov. 2002.
- Snoeij P., Attema E., Davidson M., Duesmann B., Floury N., Levrini G., Rommen B., Rosich B., "Sentinel-1 radar mission: Status and performance," *IEEE Aerospace and Electronic Systems Magazine*, vol.25, no.8, pp.32-39, Aug. 2010.
- Elefante S, Imperatore P., Zinno I., Manunta M., Mathot E., Brito F., Farres J., Lengert W., Lanari R., Casu F., "SBAS-DINSAR Time series generation on cloud computing platforms", *Proc. of IEEE IGARSS 2013*, pp. 274-277, Melbourne (AU), July 2013.

Sansosti E., Berardino P., Manunta M., Serafino F., and Fornaro G., "Geometrical SAR image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2861–2870, Oct. 2006

Pepe A. and Lanari R., "On the extension of the minimum cost flow algorithm for phase unwrapping of multitemporal differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2374–2383, Sep. 2006.

Casu F., Elefante S., Imperatore P., Zinno I., Manunta M., De Luca C., Lanari R., "SBAS-DInSAR Parallel Processing for Deformation Time Series Computation", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014

Hager G and Wellein G. *Introduction to High Performance Computing for Scientists and Engineers*, CRC Press, 2010.

El-Rewini H. and Abd-El-Barr M., *Advanced Computer Architecture and Parallel Processing*, John Wiley & Sons, Inc, 2005.

Zinno I., Imperatore P., Elefante S., Casu F., Manunta M., Mathot E., Brito F., Farres J., Lengert W., Lanari R., "A Novel Parallel Computational Framework for Processing Large INSAR Data Sets", *Living Planet Symposium 2013*, Edinburgh (GB), Sept. 9-13, 2013.

M. Bonano, M. Manunta, A. Pepe, L. Paglia and R. Lanari, "From Previous C-Band to New X-Band SAR Systems: Assessment of the DInSAR Mapping Improvement for Deformation Time-Series Retrieval in Urban Areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 1973–1984, Apr. 2013.

M. Bonano, M. Manunta, M. Marsella, and R. Lanari, "Long-term ERS/ENVISAT deformation time-series generation at full spatial resolution via the extended SBAS technique," *Int. J. Remote Sens.*, vol. Vol. 33, no. February 2012, pp. 4756–4783, 2012