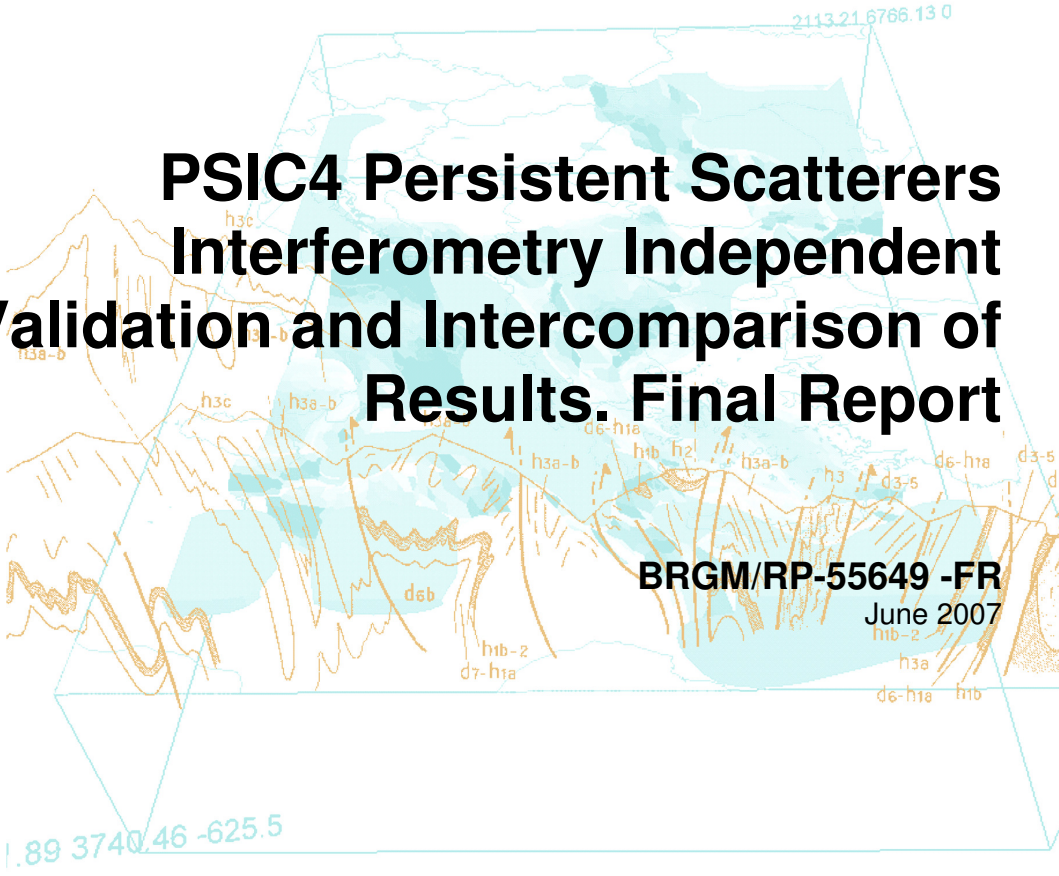
PSIC4 Persistent Scatterers Interferometry Independent Validation and Intercomparison of Results. Final Report



BRGM/RP-55649 -FR
June 2007

Geoscience for a sustainable Earth

PSIC4

Persistent Scatterer Interferometry Independent Validation and Intercomparison of Results.

Final report

PSIC4 validation Team

Daniel Raucoules, Bernard Bourgine, Marcello de Michele, Gonéri Le Cozanet, Luc Closset (BRGM)

Chris Bremmer, Hans Veldkamp (TNO)

Douglas Tragheim, Luke Bateson (BGS)

Michele Crosetto, Marta Agudo (IG)

and

Marcus Engdahl (ESA)

Reference documents

PSIC-4 Validation Site & Reference Data Description
PSIC-4 Deformation Field Characterisation Report
PSIC-4 Validation and Intercomparison Plan
PSIC-4 Conventional DInSAR processing and accuracy assessment over the validation site

Keywords: PS Interferometry, Coal Mining, Gardanne, levelling

In bibliography, this report should be cited as follows:

Raucoules D., Bourgine B., de Michele M., Le Cozanet G., Closset L., Bremmer C., Veldkamp, H., Tragheim D., Bateson L., Crosetto M., Agudo M., 2007, Persistent Scatterers Interferometry Independent Validation and Intercomparison of Results – Final Report, BRGM/RP-55649-FR

CONTENT

1. Introduction

1.1 Radar interferometry and PSI

The technique of repeat pass space-borne interferometry - that uses the phase of the signal measured by Synthetic Aperture Radar (SAR) onboard a satellite - was first demonstrated with data from the brief SEASAT mission in 1979, and later with the Shuttle Imaging Radar Missions.
The principle is based on the computation of the difference (or *interferograms*) of phase images acquired at different dates. Interferograms contain geodetic information that can be used - depending on the conditions of acquisition - for deriving Digital Elevation Models of measuring displacement of the surface of the ground.
It allows to detect and to assess under certain conditions ground deformation of small amplitude (centimetric deformations) with a good spatial resolution (decametric) and offers a synoptic view of the phenomena (ground subsidence, landslides, seismic deformation,…) over extended areas (the size of the images is about 100km, depending on the sensor).
With the sustained operation of the ESA ERS (followed by the Envisat satellite) missions over a period of more than 10 years, this technique has been validated and has delivered significant new scientific results in very diverse domains ranging from volcanology to glaciology. A major interest of these techniques resulted from the existence of this permanently increasing global archive providing the possibility of past (since 1991) deformations even without pre-existing ground-based monitoring tools. That makes SAR interferometry and derived techniques very unique among the geodetic tools.
In the late 1990's it was discovered that some bright radar targets retain their phase and amplitude stability for a period of months or years (*i.e.* on a big number of radar images) and the phase information of this targets (here denoted Persistent Scatterers – PS) can be exploited even on otherwise uncorrelated long time scale interferograms.
This led to the development of innovative methodologies that enable the detection of relative position of the PS in the direction of the satellite line of sight (LOS) with millimetric precision. These methodologies are collectively referred to as PSI-methods, where PSI stands for Persistent Scatterers Interferometry.
With the PSI techniques it is possible to exploit the phase information present in large stacks of images and to reconstruct time series of relative PS positions on areas covered by ERS1-2 and Envisat SAR images archives (since 1992). This can provide unique and indispensable information in many application fields such as ground subsidence detection, monitoring of potentially unstable structures and landslide monitoring. For summarizing, the PSI methods are a group of technologically advanced methods for measuring relative positions of ground natural targets.

1.2 PSIC4 Project

People outside the radar community have little or no experience in utilizing the products delivered with PSI methods so some concerns about PS data accuracy and trustworthiness arose amongst end-users. This issue was clearly identified during the 2003 FRINGE Workshop. ESA then decided to initiate a validation project, named PSIC4, in order to assess the performances of PSI for land deformation monitoring. In the project participated eight PSI teams, whose results have been than analysed by an independent validation team. These are some of the key questions of the project:

- How well PSI describes land deformation, spatially and temporally?

- How accurately, and how precisely PSI describes land deformation? These questions are addressed by the validation activities, where the results of each team are compared against reference data.

- How consistent are the PSI results between different teams? This question is addressed by the inter-comparison activities, where the results of the teams are compared to each other.

This document describes the work performed during PSIC4 validation and intercomparison activities. In the project, six of the eight PSI teams of PSIC4 used different implementation of the Persistent Scatterers techniques, while other two teams used slightly different concepts (basically two coherent-based approaches). In this document the eight teams are named "PSI teams". These teams have been asked to process the SAR data according to their own PSI methodology without having been given *a priori* information about the type and location of deformation in the chosen test area. The test area was selected by ESA among those proposed by the validation group considering the characteristics of the deformation (size, rate) and the levelling data availability. A consortium of independent entities has been chosen by ESA and put charge of developing and applying a methodology for intercomparison of PSI results and the validation against 'ground truth' levelling data.

The PSIC4 activities were conducted in two phases. The first phase consists of PSI processing over the validation test site. This Phase is referred to as Phase I and has been independently conducted by the 8 radar-specialist teams. The work referred to as Phase II consists on the validation and intercomparison of PSI Phase I results. The detected movements are compared against external validation benchmarks such as levelling points (validation activities) and among each other (intercomparison activities). Phase II, has been conducted by an independent consortium consisting of geological survey organizations (BRGM, BGS, TNO) and an organization in specialised in geomatics (IC). The validation and intercomparison has been performed on an anonymous basis, each team having been assigned a number by ESA. The data to be evaluated was provided uniquely by ESA without direct contact with the tested teams; so the consortium did not know which team has produced which PS results. In this report we go through the different steps of the project and show the most significant results of the validation and intercomparison activities.

Besides Phases I and II there were other complementary activities, whose results are collected in this final report. These activities are briefly described below.

1) Firstly, as already mentioned above, Phases I and II generated the main outputs of the PSIC4 project. In this context it is important to underline that the validation and inter-comparison activities of PSIC4 did not concern the specific PSI methodology used by each team. In fact, they were only focused on the outputs provided by the teams, checking how they compare with independent reference data.

2) Starting from the results achieved in the validation and inter-comparison analyses, the PSIC4 validation team prepared a set of "Open questions" to be discussed during the PSIC4 Final Presentation and the PSI validation workshop, held in ESRIN the 18-19th September 2006. These questions were organized in two main categories: those directly related to the PSIC4 results, and more general questions related to the PSI techniques.

3) The responses to the "Open questions" collected during the PSIC4 Final Presentation and the PSI validation workshop, with additional contributions collected after these two events are collected in the Appendix A of this Final Report. This document represents a valuable part of the PSIC4 project.

4) Finally, it is worth mentioning that within another ESA-founded project, TERRAFIRMA, an additional inter-comparison study was performed on the same dataset of PSIC4. This work only concerned four of the eight teams of PSIC4. The final report of this study, named "Provence Inter-comparison", is reported in the Appendix B of this report.

2. Characterisation of the validation site and levelling dataset

This chapter describes and characterises the test site selected for the PSIC-4 project and the available ground data sets to be used in the second phase of the project.

The selected area of interest is located around the town of Gardanne in the sedimentary basin of “l’Arc” between the Sainte Victoire mountain, the cities of Marseilles and Aix en Provence. The area of Marseilles-Aix is the second biggest urbanized area in France (more than 1.4M inhabitants).

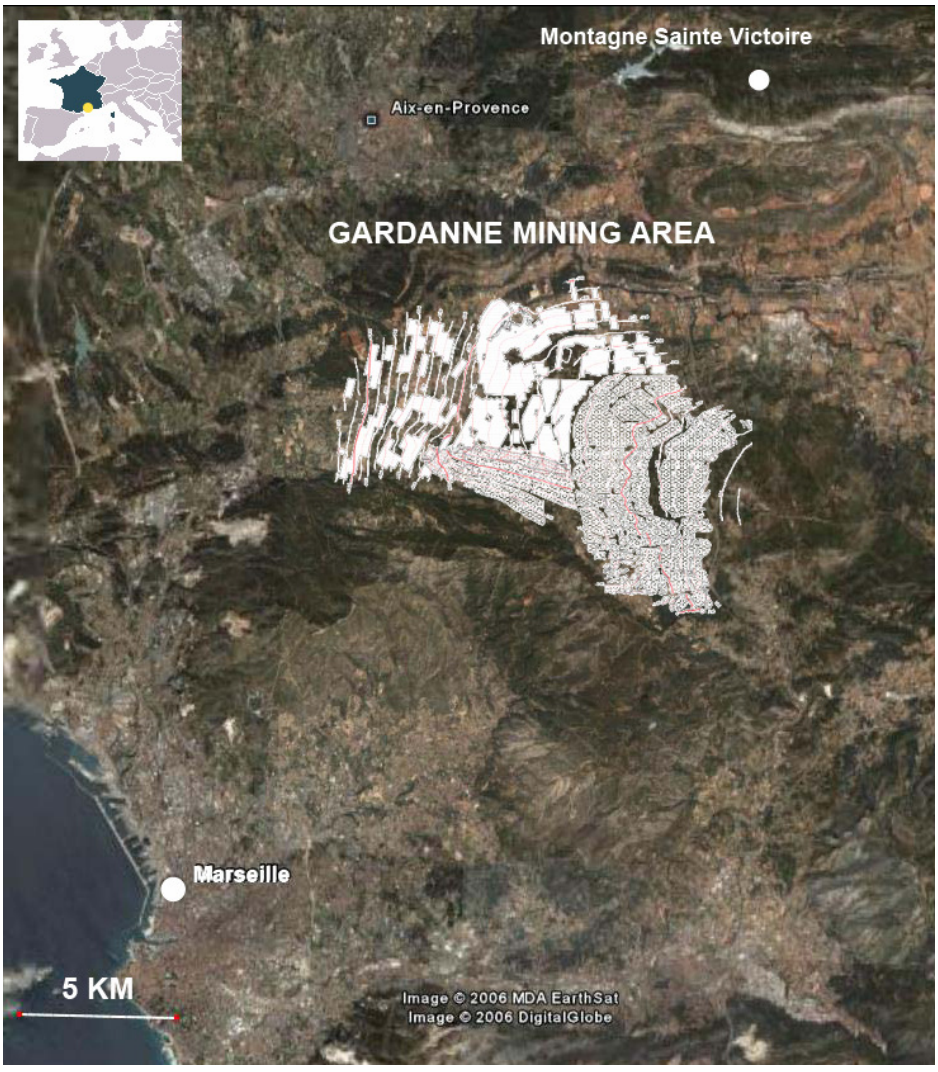


Figure 2.1 The Gardanne mining area

Gardanne coal deposit has developed in a particular geologic context called Basin of Arc. During the Santonian period (76My), the sea in the lower Provence Gulf started to regress. After this event, rivers and lakes dominating the landscape, sedimentary cycles and accumulation of vegetal debris generated the coal layer in this new fluvio-lacustral basin. The Coal deposited in carbonated rocks (limestone) and represents one of the most recent coal deposits in France, in contrast to other French coal deposits belonging to Carbonifer (280My) The basin is composed by carbonates and clay formations with a thickness of about 1500m. Structurally speaking the Basin of Arc is a synclinal fold oriented East-West and limited by folded structures and strike-slip structures. In particular, the Sainte-Victoire and Etoile structure.

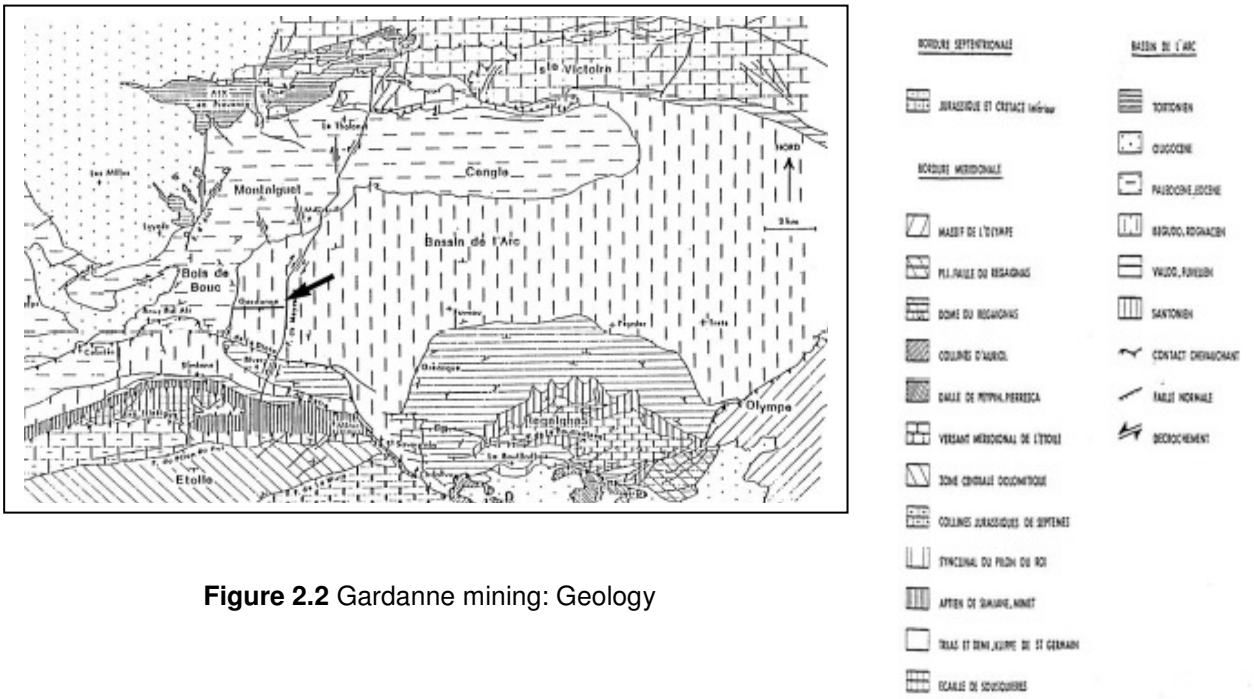


Figure 2.2 Gardanne mining: Geology

The coal (lignite) exploitation on the area started during Middle Age but the structured industrial production began during century XIX. The maximum of production corresponds to the 1965-1993 period. The exploitation of the concession stopped on February 1 2003.
Different mining techniques were used on the area. By using the *room and pillar* technique, pillars are left in place in a regular pattern while the rooms are mined out. The pillars avoid the collapse of the roof during the coal extraction. With the *long wall mining technique*, a long wall (about 250 m long typically) of coal is mined in a single slice (typically 1-2m thick). The longwall "panel" is typically some km long and 250m wide. The longwall equipment consists of a number of hydraulic jacks, which are placed in a long line up to 400m in length in order to support the roof. The coal is cut by a shearer that moves along the length of the face extracting the coal. As the shearer removes the coal, the equipment move forward into the newly created cavity. As mining progresses, the cavity behind the longwall collapses under the weight of the overlying materials.

During the period of interest (1991-2004), the observed ground subsidence is associated to the coal mining exploitation using long wall mining technique. This technique has particular consequences in terms of size and evolution (fast evolution) of the deformation compared to techniques (chamber and pillar) used in the older exploitations of the area.
In fact, typically, when a panel is mined, we expect that half of the displacement is produced in the first 2 months and the residual deformation during the 3 following years. A point in the surface can be influenced by several mined panels and therefore the deformation may last on longer periods. This behaviour is examined in section 3.

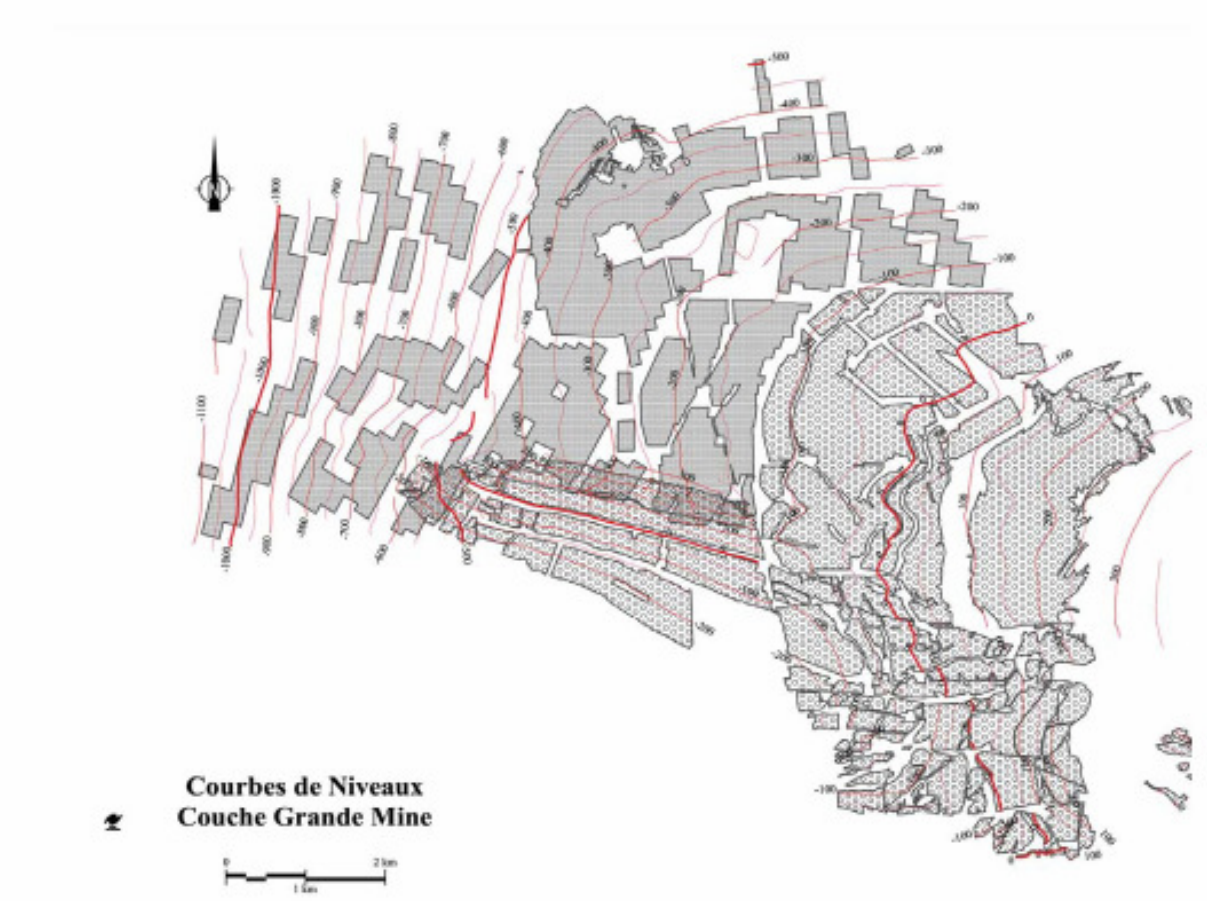


Figure 2.3: Detailed plan of the underground mine panels. Isolines indicating the depth of the exploitation panel are in red.

The monitoring of the land deformation effects associated with the coal mining exploitation of the Gardanne area is performed through spirit levelling surveys. The Topographic Service of the Gardanne Mine has always based its surveys on precise levelling tools.
The in situ data of the Gardanne test site have been acquired in the last decades in order to monitor the deformation effects on the surface associated with the coal mining exploitation using long wall mining technique, which is realized in the Gardanne area. The Topographic Service of the Gardanne Mine (Service Géomètres) is responsible for the deformation measurements of the area, which first started in June 1972 for the monitoring of a channel of the "Société du Canal de Provence" and which are still regularly performed. Since then the monitoring network has been considerably expanded in order to include other features, e.g.:

- the water column of Rousset, which feeds a thermal plant, started in July 1982,
- the railways line Aix – Marseille of SNCF, started in January 1986,
- the factory Pechiney, started in August 1986,
- the tower Puits Y, started in July 1990,
- the village of Bouc Bel-air, started in July 1990,
- the village of Simiane, started in April 1993,
- the city of Gardanne, started in October 1994.

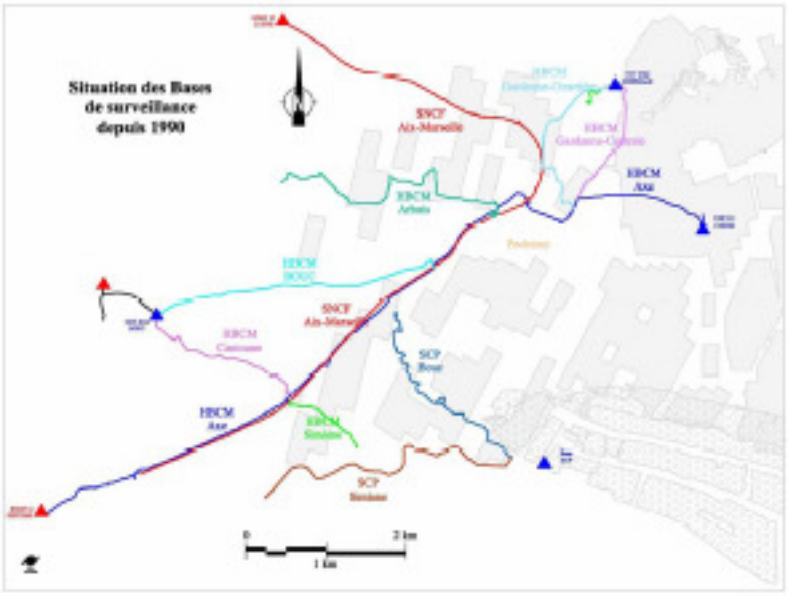


Figure 2.3: Gardanne mining area, levelling network

Since 1990 a new levelling instrument is used, the Wild NA3003, which is an automatic electronic level with bar code. Again, this is an excellent tool for precise levelling measurements, with the following theoretical precision performances:

Standard deviation of ΔH, measured point-to-point = ± 0.7 mm
Standard deviation of 1 km one way levelling = ± 1.5 mm

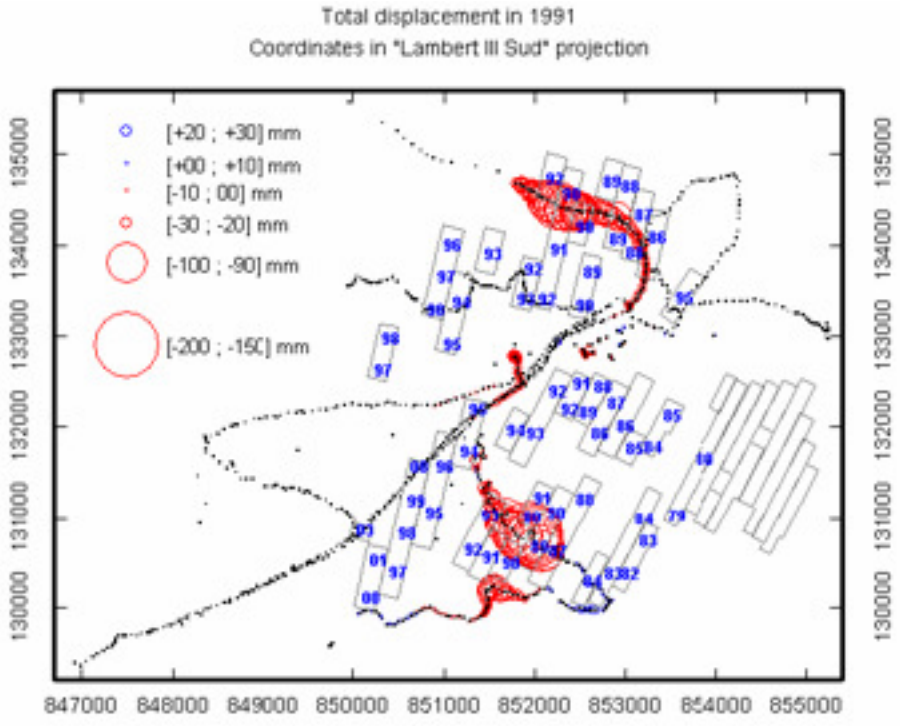
In that concerns the validation data, the spatial and temporal coverage of the levelling is sufficient to carry out the comparison. The accuracy of the levelling is consistent with the expected accuracy of the PSINSAR at given dates (3-5 mm). We can therefore assume that the characteristics of the ground data are fully acceptable for the project.

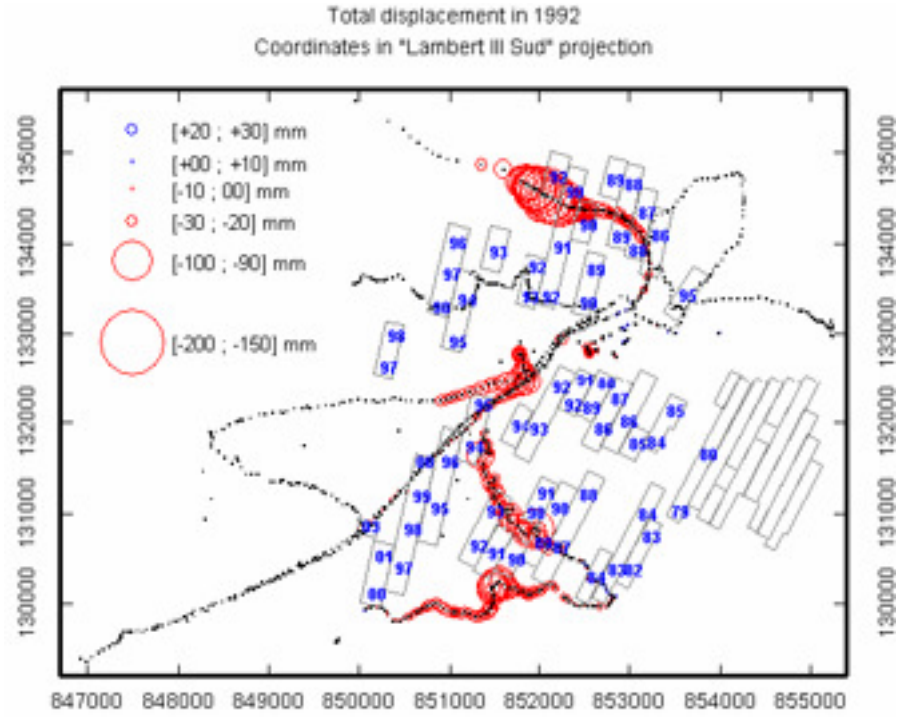
3. Characterisation of the deformation field

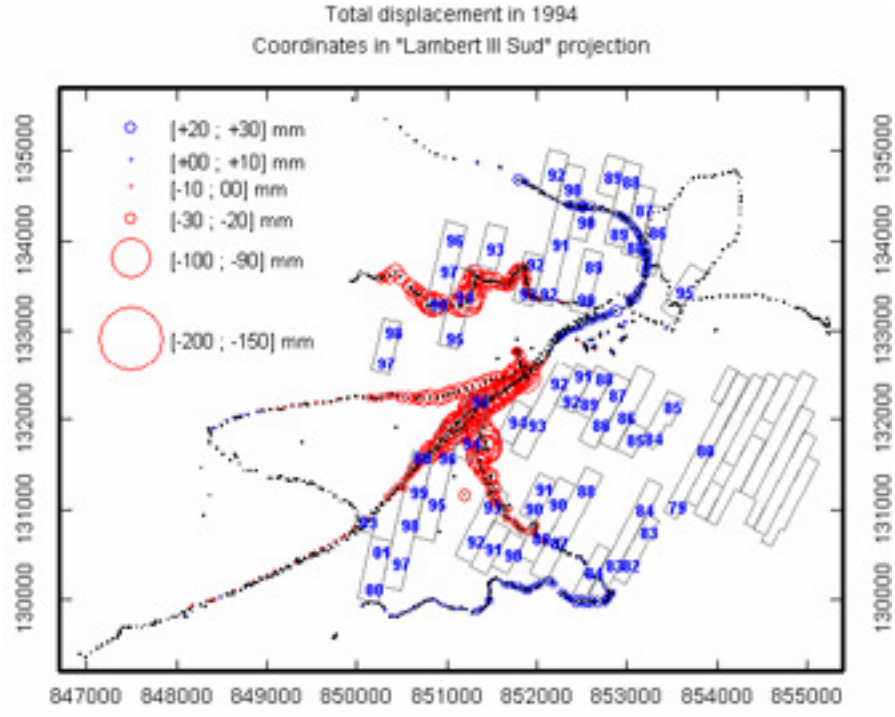
The characterisation of the deformation field is based on the analysis of *Charbonnages de France* levelling data. Levelling data consists of 12 survey lines along a railway line, along canals, etc. with an inter-distance between points of 40 to 50 m and a measurement of elevation at more or less regular time intervals (every 6 months in average).
The surveyed area extent is about 8 km x 6 km. The figure below shows the location of the survey lines. Mining works (longwall) are also displayed on this figure.
During the period of interest (1992-2004), the observed ground subsidence is associated to the coal mining exploitation using long wall mining technique. This technique has particular consequences in terms of size and evolution (fast evolution) of the deformation if compared to techniques (chamber and pillar) used in the older exploitation areas.
Typically, for the concerned period, the exploited panels are located between **600m** and **1100m** depth (appendix 2) with width of about **250m**.The thickness of the exploited coal layer is between 2 and 3 m. The positions of the panels affecting the area during the monitored by PSInSAR are located on figure above.
The production (stopped in 2003) was about 1,15 million ton/yr.

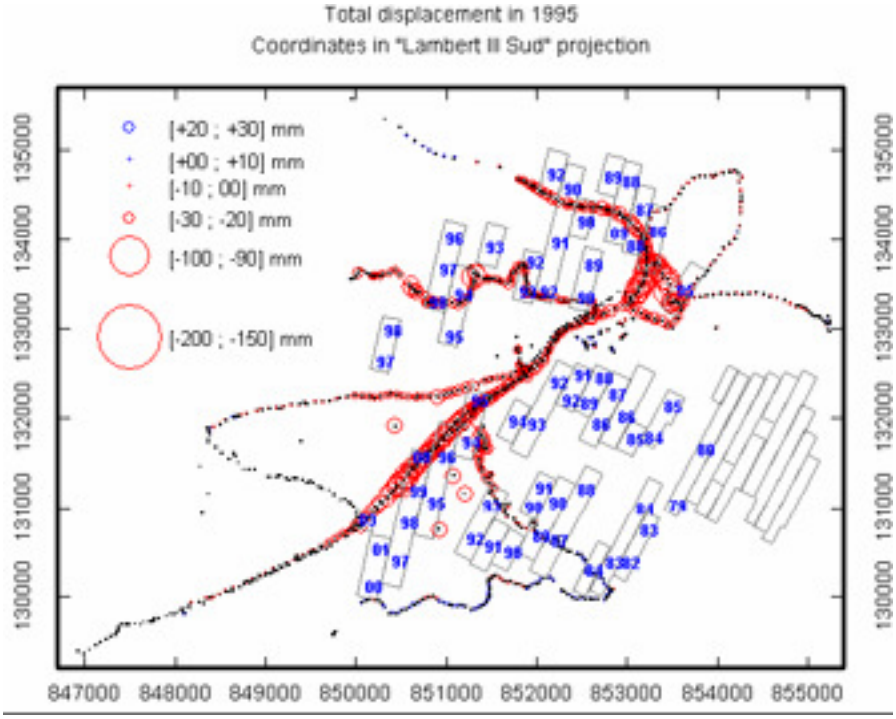
3.1 Evolution of deformation as a function of time and space

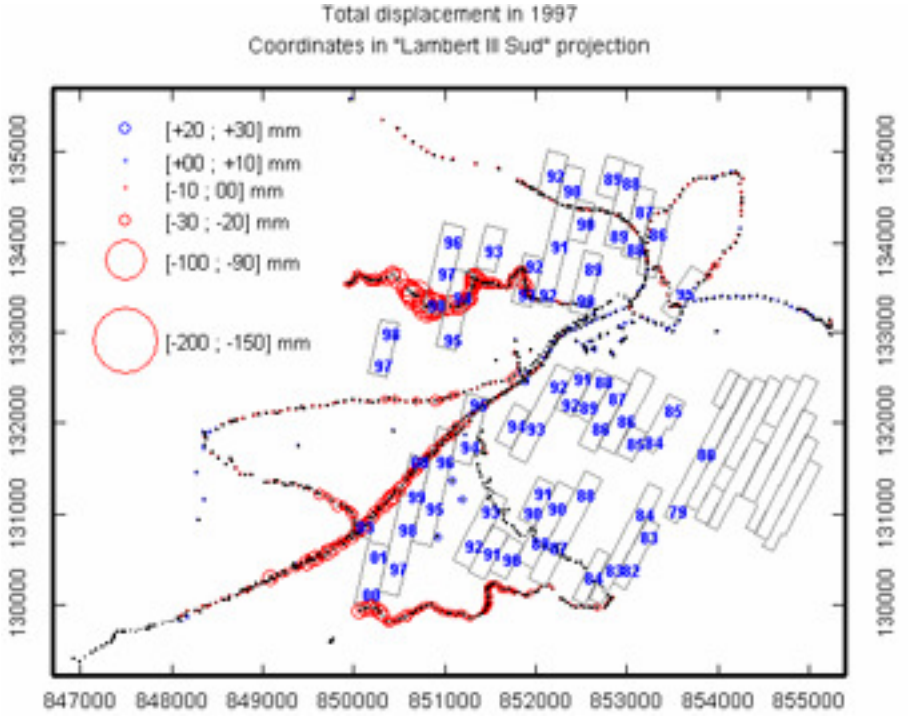
The figures below show the annual displacement from 1991 to 2001, with symbols proportional to the displacement. The maximum displacement observed within a year if of about 170 mm, and occurs in 1991. In average, taking into account the levelling points that remain stable, the average annual displacement is always lower than 5 mm.
For a given point, 70 % of the subsidence occurs within the 3 first months after the arrival of mining works in the corresponding area. It is admitted that 2 years after, this point became stable. When several works overlap, their influences cumulate and the point continues to move for more than two years.

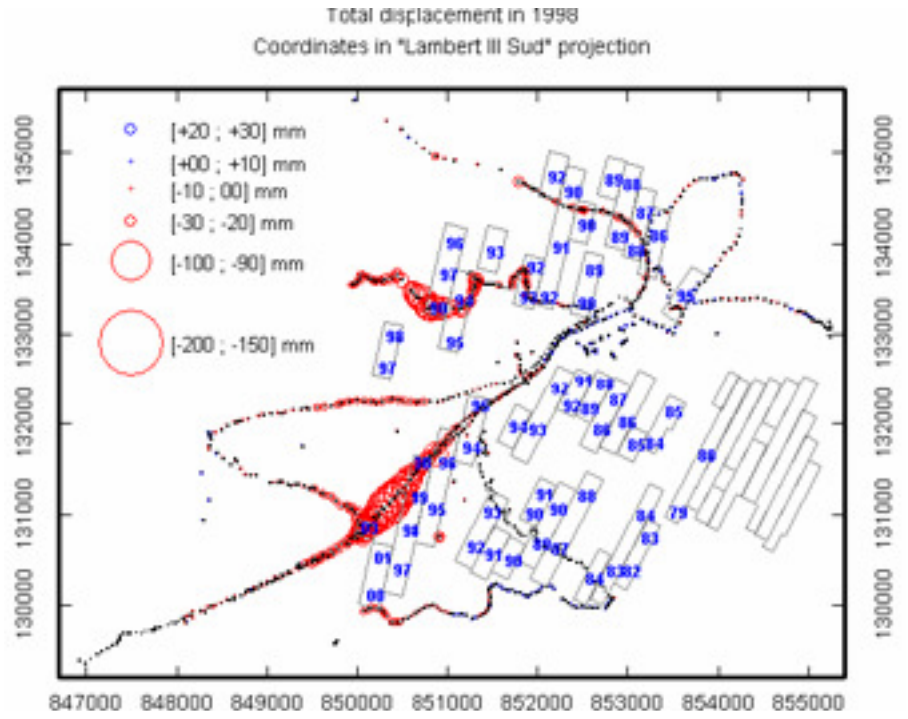

Total displacement in 1991
Coordinates in "Lambert III Sud" projection
[+20 ; +30] mm
[+00 ; +10] mm
[-10 ; 00] mm
[-30 ; -20] mm
[-100 ; -90] mm
[-200 ; -150] mm
847000 848000 849000 850000 851000 852000 853000 854000 855000
130000 131000 132000 133000 134000 135000


Total displacement in 1992
Coordinates in "Lambert III Sud" projection
[+20 ; +30] mm
[+00 ; +10] mm
[-10 ; 00] mm
[-30 ; -20] mm
[-100 ; -90] mm
[-200 ; -150] mm
847000 848000 849000 850000 851000 852000 853000 854000 855000
130000 131000 132000 133000 134000 135000


Total displacement in 1994
Coordinates in "Lambert III Sud" projection
[+20 ; +30] mm
[+00 ; +10] mm
[-10 ; 00] mm
[-30 ; -20] mm
[-100 ; -90] mm
[-200 ; -150] mm
135000
134000
133000
132000
131000
130000
847000
848000
849000
850000
851000
852000
853000
854000
855000


Total displacement in 1995
Coordinates in "Lambert III Sud" projection
[+20 ; +30] mm
[+00 ; +10] mm
[-10 ; 00] mm
[-30 ; -20] mm
[-100 ; -90] mm
[-200 ; -150] mm
135000
134000
133000
132000
131000
130000
847000
848000
849000
850000
851000
852000
853000
854000
855000


Total displacement in 1997
Coordinates in "Lambert III Sud" projection
[+20 ; +30] mm
[+00 ; +10] mm
[-10 ; 00] mm
[-30 ; -20] mm
[-100 ; -90] mm
[-200 ; -150] mm
847000 848000 849000 850000 851000 852000 853000 854000 855000
130000 131000 132000 133000 134000 135000


Total displacement in 1998
Coordinates in "Lambert III Sud" projection
[+20 ; +30] mm
[+00 ; +10] mm
[-10 ; 00] mm
[-30 ; -20] mm
[-100 ; -90] mm
[-200 ; -150] mm
847000 848000 849000 850000 851000 852000 853000 854000 855000
130000 131000 132000 133000 134000 135000

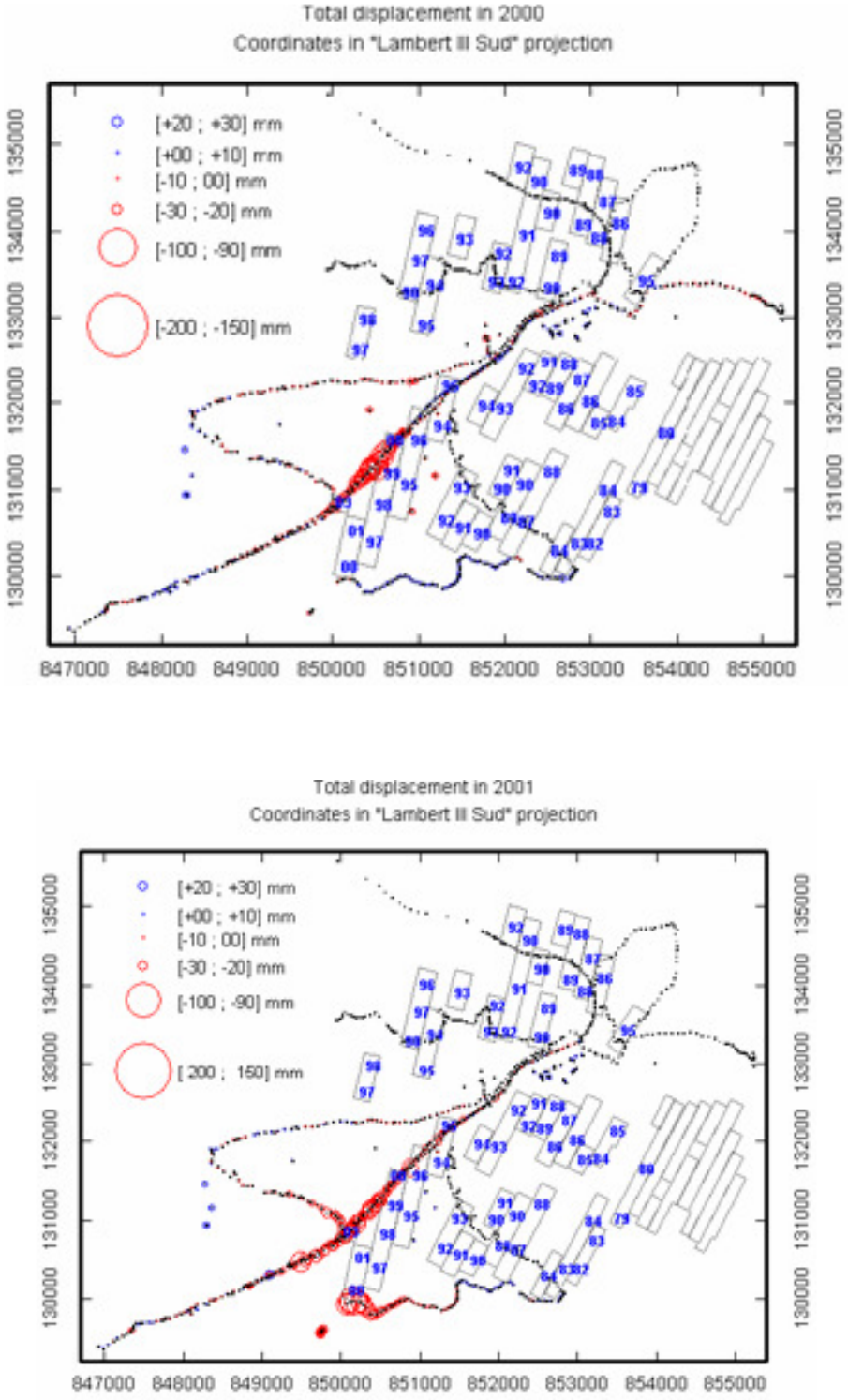


Figure 3.1: annual displacements from 1991 to 2001 observed by levelling on the area and mining map. The year of mining is indicated on the panels.

We could say that the images above represent the deforming areas at different epochs. These images give an idea about the amplitude and location in time and space of the occurring deformation on the Gardanne mining site at a certain moment in time. Besides, it is interesting to show the total amount of deformation that occurred across the whole study period 1992-2004 on the test site. We call this amount of deformation, cumulative deformation over the time period 1992-2004. In actual fact, it is this cumulative deformation (calculated by levelling values resampled at SAR acquisition dates) that we will use as

reference for comparison against deformation measured by PSI technique in the validation activity.

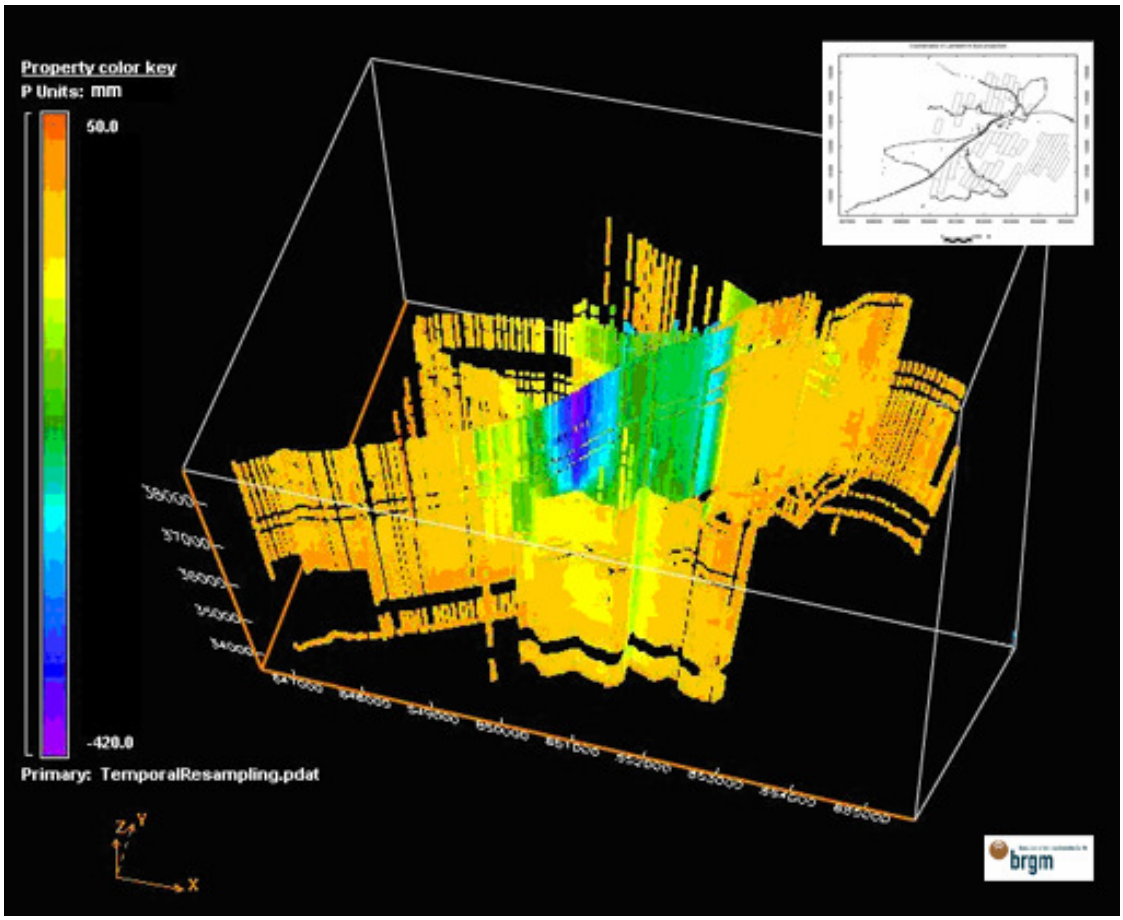


Figure 3.2 Cumulative deformation measured by levelling

4. Validation activities

Phase one of the project produced 8 set of data containing the PSI information calculated by the 8 teams. The sets of data have been delivered to the validation consortium in the form of excel spreadsheets. Each team has been given the same number of SAR scenes to process, though they have been given the freedom of choice about which data subset (in time and space) to use.
Columns in the spreadsheets contain the following values: ID of the point; Latitude; Longitude; Ellipsoidal Height; Range radar coordinates; Azimuth radar coordinates, Line Of Sight Velocity; Phase Coherence; Time Series of PS heights; Time Series of Atmospheric Phase Screen values. All the teams provided the aforementioned data, though team 5 did not provide the consortium with the APS values.
The teams have been given an Orthophoto of the Gardanne area (0.5 metres spatial resolution) in order for them to be able to adjust and ameliorate geocoding of PS, if and where corrections were thought needed.

4.1 Pre-validation steps

The validation consortium analysed the 8 data files and encounters some discrepancies and inhomogeneities between the files. For instance, none of the teams found the same number of PS, PS density changes from one team to another one, none of the teams used the same subset of the SAR frame. The aforementioned differences might depend on the different process and work flow adopted by each team. Besides, some other technical discrepancies have been encountered. Firstly, most of the teams had performed a misleading geocoding, which did not allow one to refer a particular PS to a physical structure on the ground (team1 and team8 in particular). In some other cases, geocoding has been performed in such a way that PS were translated in X and Y if compared to features in the orthophoto (teams2, 3, 4, 5, 6 and 7). Moreover, none all the team had selected the same reference SAR date (the date chosen to represent the start of deformation analyses). Therefore, PS data could not be validated and compared straight forward so that the consortium had to perform 4 'pre-validation' adjustments.

The following pre-validation steps were performed:

Coordinate transformation (task responsible, BRGM)

The Lambert III Sud map projection is the conventional map projection for South of France and the extended Lambert II a map projection for the whole Country, the orthophoto and the levelling data were provided by IGN and CDF respectively in extended Lambert II in Lambert III map projection. In order for PS data to be comparable to levelling data, the two quantities are supposed to be in the same reference system. Therefore, the validation consortium converted PS coordinates from geographic latitude and longitude into the Lambert III map projection.

Correcting for geocoding errors (task responsible, BRGM)

PS data presented a shift in the geo-position if compared to ground features in the orthophoto. Each team presents different degrees of mismatching between ground features and PS location. For instance, team 1 and team 8 seemed not to have any physical relation between ground features and PS positions. For these two teams we did not apply any

geocoding compensation while all the other teams have been compensated by linear shifting the PS data according to a well identified scatterer on the ground. The table below show the total amount of geocoding compensation needed to be applied to the teams.

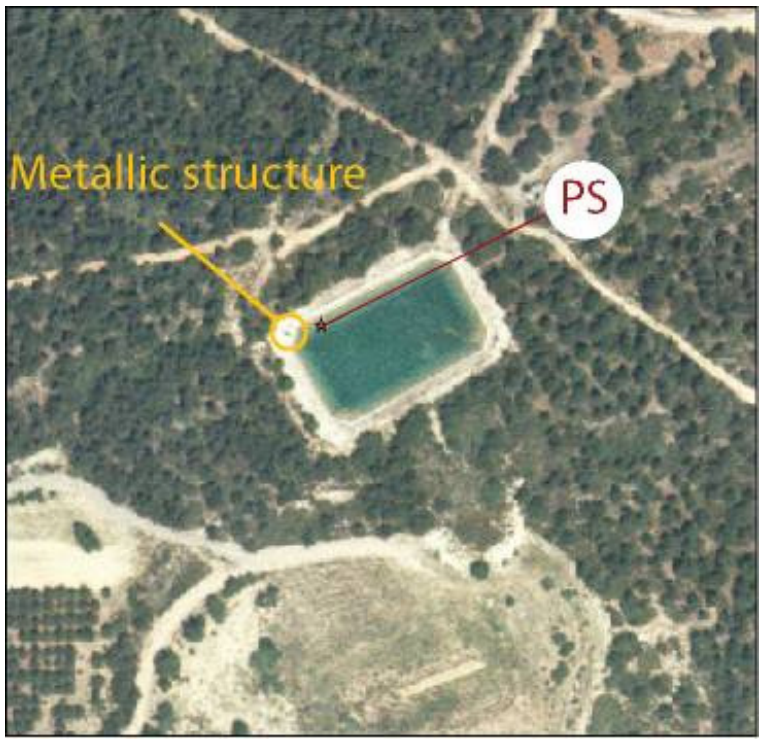


Figure 4.1. This figure illustrates the geocoding bias (example of Team 4). Since this PS unambiguously refers to the metallic structure at the side of the pond, we could use the distance between them to compensate the geocoding error.

TEAM	dx	dy
1	No reliable correspondence	
2	15.35	-2.21
3	4.84	2.38
4	13.88	4
5	9.46	-1.47
6	-79.36	21.12
7	10.51	-10.62
8	No reliable correspondence	

Referring PS deformation to the same date and reference area *(task responsible, BRGM)*

The PSIC4 data processing has been performed by the 8 teams in the frame of a blind test. It means that the teams did not have any information *a priori* regarding the spatial extension, the amount and the time occurrence of the deformation. As the teams used different reference stable areas, the consortium decided to refer the PS dataset to the same reference stable area in order for the PS data to be comparable. As well, the PS datasets have been referred to the same 'zero deformation' common date. We selected a stable area according to the levelling measures. There are 5 levelling benchmarks within the chosen stable area.

The measurement on those 5 levelling points over the time period 1992/2004 yields zero (±3mm) deformation. We corrected each PS LOS velocity and Time Series by subtracting the mean PS Line Of Sight (LOS) velocity on the area to all PS LOS velocity. In this way both levelling and PS values yields 'zero' within the stable area. The reference date (15/07/1992) represents the earlier SAR dataset that all the teams have in common and so it could be chosen as 'zero deformation' reference date.

Team number	N° of PS within	Mean mm/y	STDV mm/y
T1	3	-0.6	0.2
T2	16	-0.8	0.8
T3	25	0.8	0.3
T4	34	0.7	0.6
T5	25	-1.2	0.8
T6	37	-1.2	0.3
T7	62	-0.7	0.5
T8	81	-0.5	0.5

Table 4.1 Number of PSs, mean and standard deviation of LOS velocities of PSs in the stable area before correction.

At this point the PS dataset are set for intercomparison though they are not set for validation. In fact, the validation process needs each PS to be exactly localised on a levelling benchmark. Moreover, the time sampling of each levelling measure should be consistent with the time sampling of PS measures. To overcome this obstacle, both levelling and PS data have been *temporally* resampled and *spatially* resampled respectively.

Spatial resampling of PS *(task responsible, TNO)*

Spatial resampling of the PS-data has been performed for each team. The rationale behind it is the fact that the spatial sampling of the PS-field and the levelling lines is incongruent. Secondly, PS-InSAR results in a massive oversampling with respect to coventional levelling. This results in the fact that in the direct neighbourhood of a levelling location a number of permanent scatterers can be found. In order to compare both results an estimation of the displacement on identical locations is necessary. By a spatial resampling of the PS-data, a co-located estimation can be made of the displacement in time at a number of levelling points and thus making it possible to compare time-series of displacement.

The corrected results of the teams (correction for geocoding error, geocoded in a Lambert III-South reference system, reference to the same stable area) have been taken as input.

For this task the geostatistical software packages Isatis version 5 (Geovariances) is used for a geostatistical analysis, and GoCAD version 2.1.2 (Earth Decision Sciences) for visualisation.

Procedure:

1) For each team a normal score transform of the cumulative displacement has been computed for each image. The semi-variograms for these normal score transforms have been visually inspected. It was concluded that these semi-variograms are stable and that, per team, a single semi-variogram can be used to model all images of the normal score transform of the cumulative displacement.

2) For each team, based on the semi-variogram at hand, an ordinary kriging computation has been performed of each image. A 50 meter search neighbourhood has been used in order to meet short distance variability. No threshold of numbers of available PS-points is applied.
3) The kriged maps are back-transformed, including the estimated cumulative displacement + or – one kriging error. The back transformed values are stored in a csv-file containing the following output.

OUTPUT: For each team a database is filled with selected interpolated PS-displacements corresponding to co-located levelling locations. This is a CSV-file consisting of the following columns:

SN+ = Sample rank number for the estimated levelling points

Xl3s = x-coördinate of levelling point in Lambert3-Sud projection

Yl3s = y-coördinate of levelling point in Lambert3-Sud projection

Day = Julian day

BREF = code assigned 1 for levelling time series starting before 15 july 1992 and assigned 0 for levelling time series beginning later than 15 july 1992

Date = date

Disp = displacement estimated by temporal interpolation of the levelling time series

Indc = levelling location, code

Type = name of levelling series

n-team = number of teams for which a displacement could be estimated at the specified location

*t*_cumulative_displacement_estimate* = estimated cumulative displacement for team * at date *Date* for levelling location *Indc*

*t*_cumulative_displacement_stdev* = kriging error for team * at date *Date* for levelling location *Indc*

The average RMSE for all individual PS-time series with respect to the levelling time series has been computed in order to asses the error associated with averaging the PS-time series over a 50-metre neighbourhood. Figure 4.2 shows the average RMSE with respect to the distance to the levelling locations. The results are based on a subset of the levelling time series. Only those time series have been chosen which have a maximum displacement inbetween SAR-acquisition dates smaller than 1,4 cm in order to have a sub-set which has no errors due to phase-unwrapping. It can be concluded that within a 50 metre neighbourhood there is no trend in RMSE with distance to the levelling point and that the kriging approach for estimating PS-time series at the levelling location is acceptable.

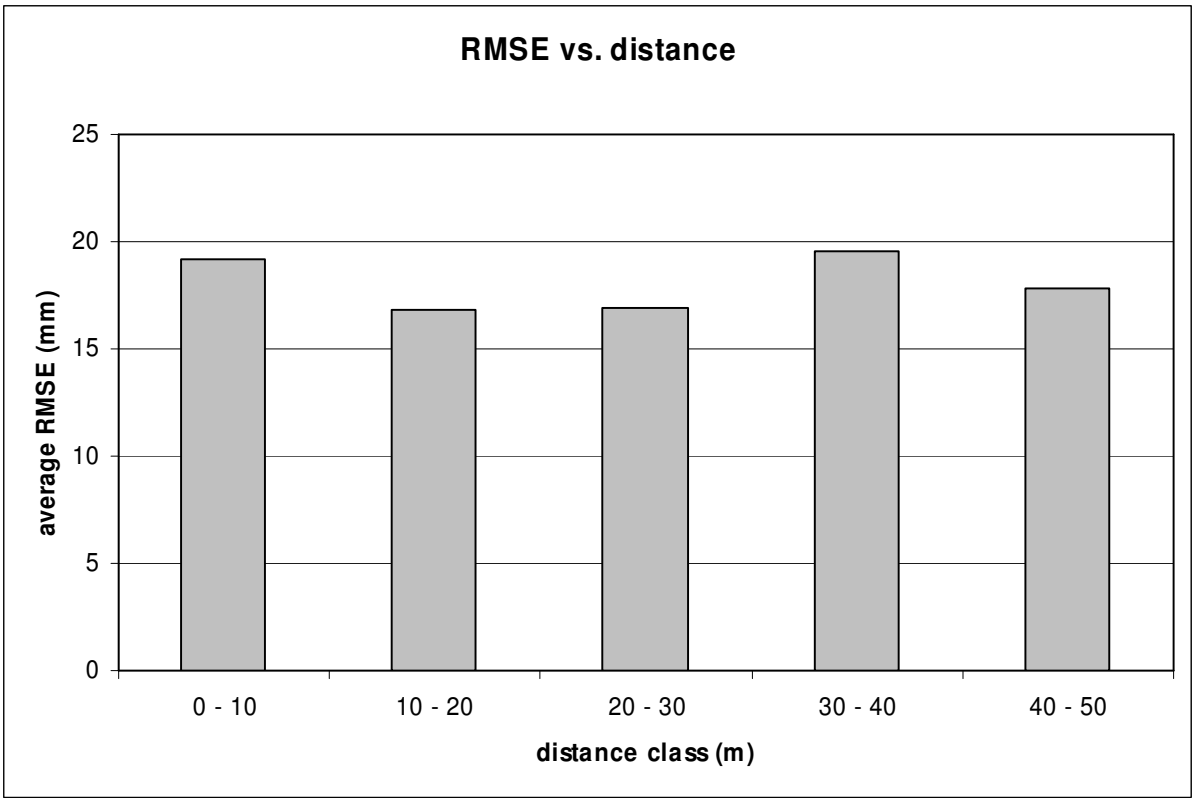


Figure 4.2.. Average RMSE (in mm) of PS-time series with respect to the levelling time series as a function of the distance (in m).

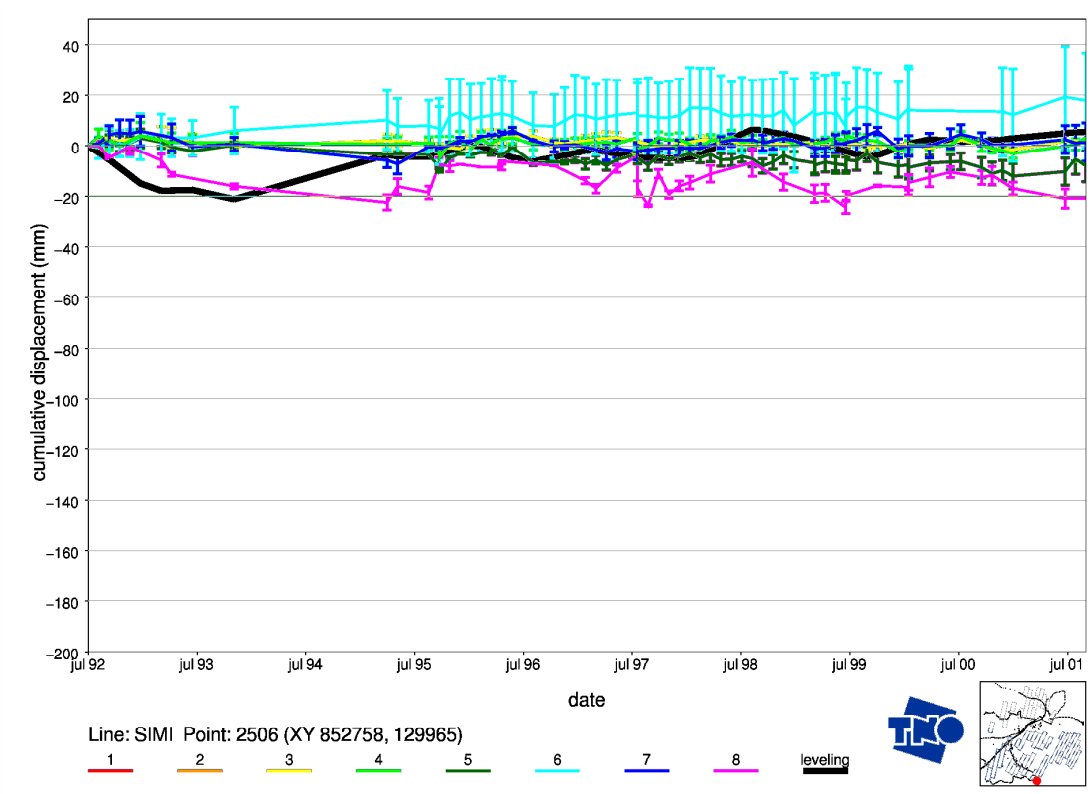


Figure 4.3 Example of the estimated PS-time series of displacements of a number of teams against the levelling time serie at a certain levelling location (in this case SIMI-2506). The error-bars show the kriging error associated to the ordinary kriging system used to estimate the PS-time series at the given location.

Temporal resampling of levelling data (task responsible, BRGM)

Levelling values have not been sampled at the same time as SAR data. In fact, SAR data over a particular area are available at least every 35 days, corresponding to the repeat cycle of the satellite, while levelling points have been surveyed every 6 months. Therefore, data acquisition do not overlap yielding a temporal shift between PS and levelling data acquisition in terms of days. This reality, does not allow one to make absolute comparisons. For this reasons we have used the variogram time model to analyse the temporal evolution of Levelling measures and then we used this information as input in a kriging resampling procedure. By this method we are able to predict levelling deformation values every 35 days instead of ~180 and exactly on the same days as SAR acquisitions. The error on the prediction tends to zero as the prediction approach the actual levelling date (see figure below). For this temporal interpolation of displacements, at a given (X,Y) location, only the measured displacements at the same (X,Y) but at different dates are used. The kriging neighbourhood is 750 days (approx. 1 year before and after the target date), so that displacements at each SAR date can be interpolated using levelling data measured before and after. Temporal resampled levelling data are set to zero on 15/07/1992 as SAR data.

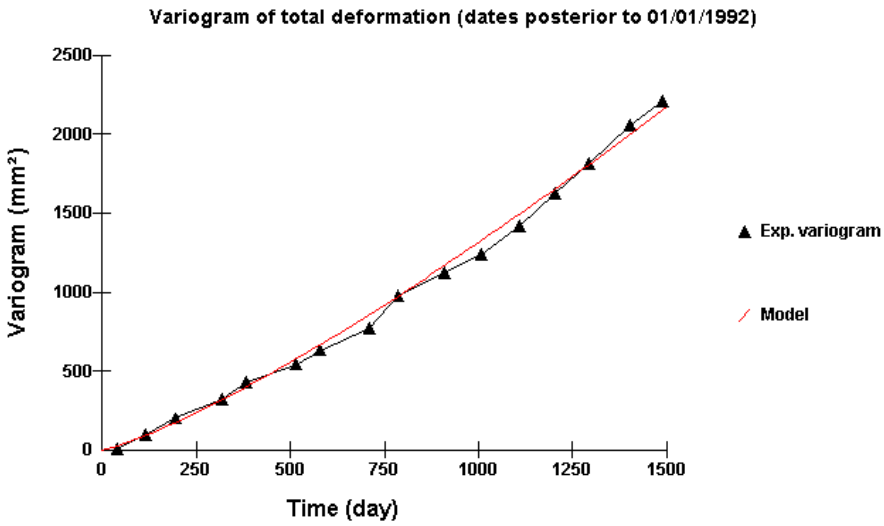


Figure 4.4: variogram of deformation measured by levelling

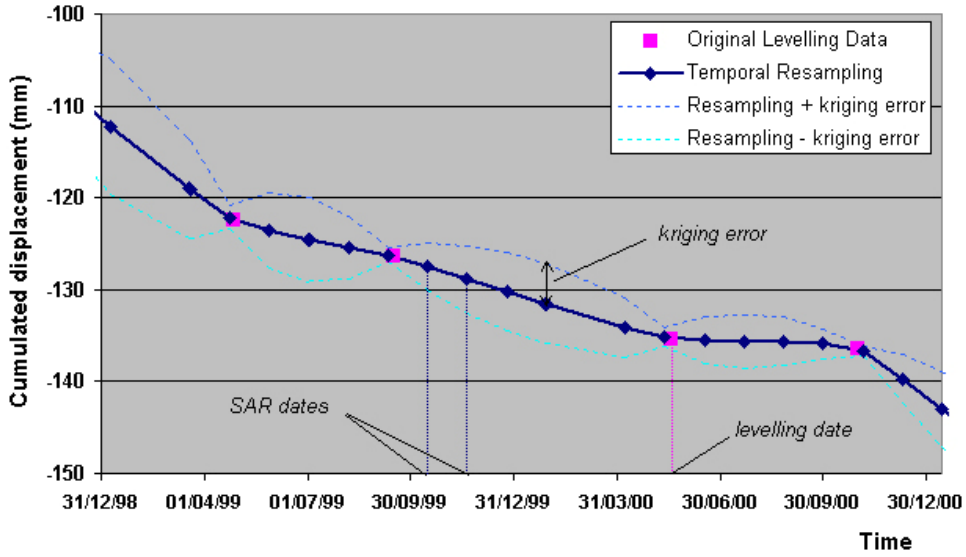
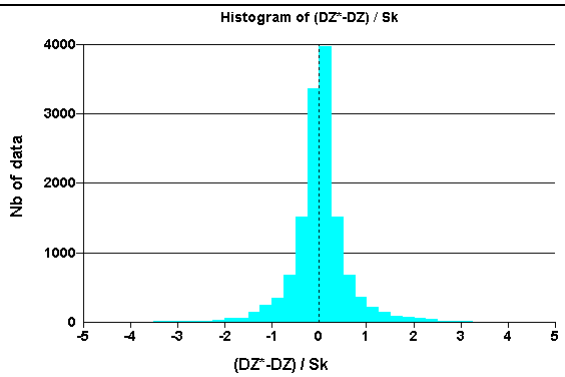


Figure 4.5: principle of the temporal resampling of the levelling (example on point RNGF 7)

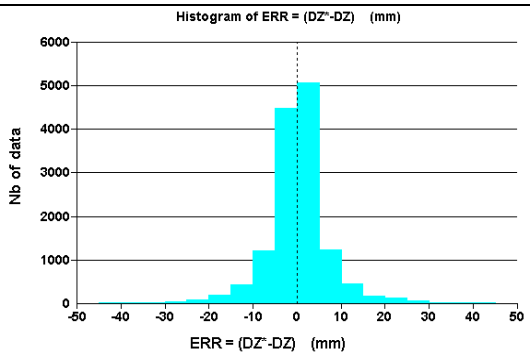
This variogram shows a temporal trend, due to the non-stationary evolution of deformation. The variogram represented here corresponds to the "average" temporal behaviour. When we compute the variogram on subsets of data, we observe that the variogram value is lower than the above average variogram when the deformation rate (the velocity) is "rather low", and higher when the deformation rate is "rather high". But the global shape remains more or less the same, *i.e.* the exponent of the power model remains close to 1.25, at least for the behaviour for time intervals lower than 1 to 2 years. In other terms there is a kind of "proportional effect" of the temporal variogram linked to the velocity and to the trend generated by the evolution of velocity. The consequence is:

- the temporal interpolation can be performed using the average variogram model: this will not affect significantly the interpolated value. It is not necessary to "detrend" the data before interpolation
- the kriging error computed from the average variogram must be corrected in order to account for the proportional effect

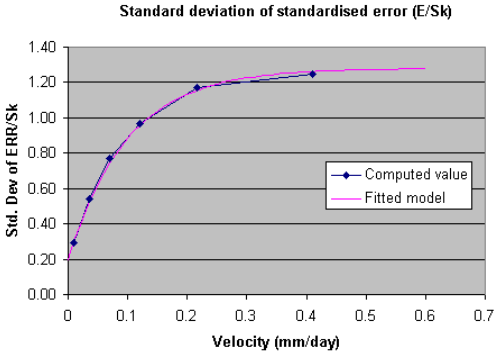
To evaluate the correction factor for the kriging error we performed a cross validation. The results of the cross validation performed on levelling data are displayed below (histogram of raw and of standardised errors). The average interpolation error is equal to -0.02 mm, showing that there is no global bias. The average standardised error is equal to 0.08 (it should be close to zero) and the standard deviation of the standardised error is equal to 0.64 (it should be close to 1)



Histogram of standardised errors

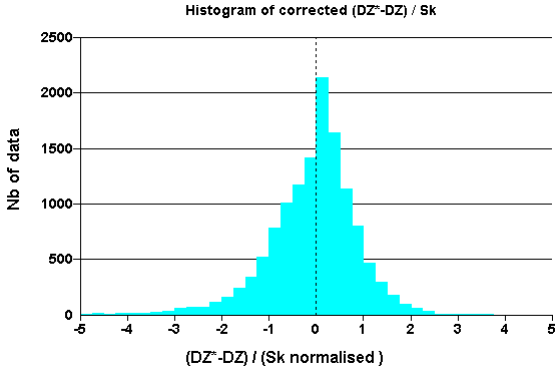


Histogram of error



Correction factor of the kriging standard deviation

Correction factor = 0.2 + 1.08 1 – Exp(-10*Velocity)]



Histogram of corrected standardised errors

Then we corrected the bias on the kriging standard deviation by evaluating the link between the instantaneous deformation rate (velocity estimated at a given date by comparing with the measurement acquired at the previous date) and the standardised error. After correction of the kriging standard deviation, the standardised errors have a mean equal to 0.1 and a standard deviation equal to 1.06. Their histogram is shown above.

The temporal resampling calculation yielded the results shown in the figure 3.2. This image shows very well the localised spatial and temporal evolution of deformation along the levelling lines.

4.2 **Check of geocoding quality (BGS)**

As already discussed in section 4.1 BRGM had to measure, and calculate a global shift in the PS points for Teams 2 to 7. This was based on the dispersion of these points (about 5) chosen reference points on an orthophoto, consisting of a pipeline intersection (846913E, 128611N), 2 pylons (854553E, 133797N and 860375E, 125678N) and 2 buildings (863948E, 127700N and 847969E, 133805N). It was not possible to relate the PS points to ground or building features for Team 1, because these points were on a widely spaced grid. In contrast, Team 8 had so many PS points, it was difficult to be sure what features the points related to. Figure 4.2.1 shows all the vectors superimposed, and which clearly demonstrates that each Teams results differed from each other in geocoding azimuth and magnitude. This was why a pre-processing shift was necessary, before any further analysis could be undertake.

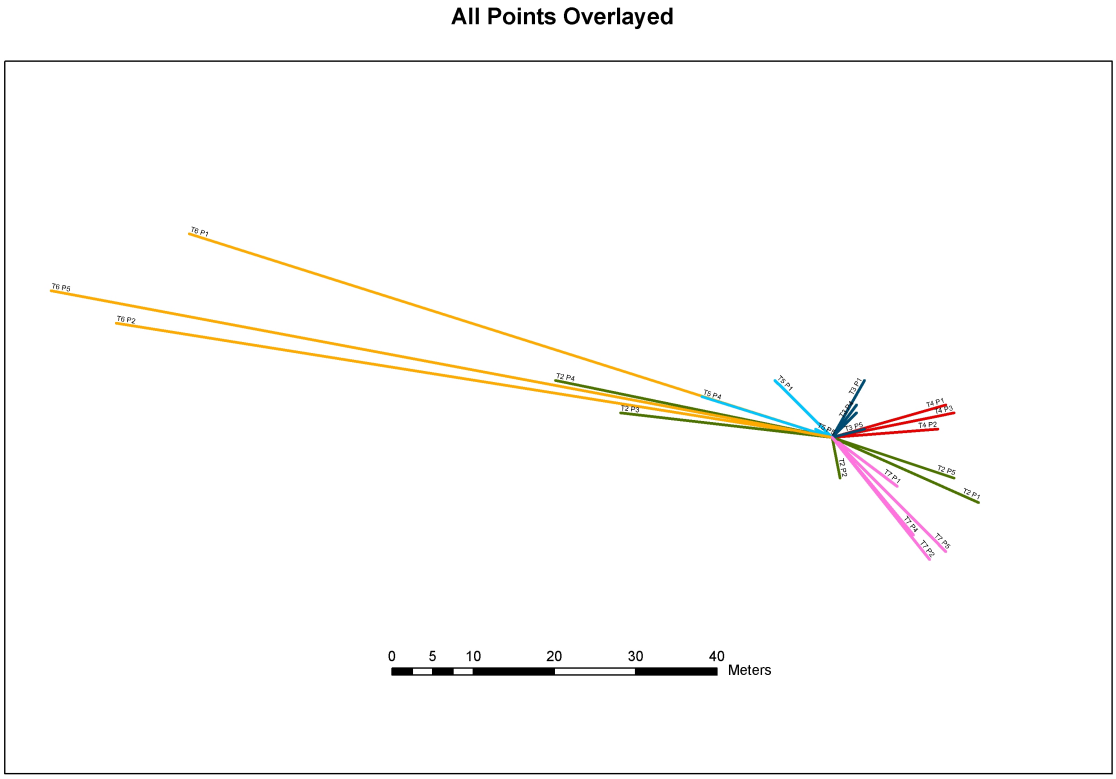


Figure 4.2.1 The original shifts from all Teams, with all vectors overlaid onto a single point origin. Team 2 in green, Team 3 in dark blue, Team 4 in red, Team 5 in light blue, Team 6 in orange, Team 7 in pink. Most shifts are within 20 meters, however Team 6 shows the largest shift distance, while Team 2 shows a mainly bimodal direction in shift.

Based on these results BRGM applied global shifts as detailed in Table 4.2.1

TEAM	dx	dy
1	No reliable correspondence	
2	15.35	-2.21
3	4.84	2.38
4	13.88	4
5	9.46	-1.47
6	-79.36	21.12
7	10.51	-10.62
8	No reliable correspondence	

Table 4.2.1. Global shifts applied by BRGM during the pre-processing stage.

Subsequent geocoding quality was assessed by BGS using the BRGM 'pre-processed' PS results.

Method:

The following account describes the step-by-step procedure as to how the geocoding of the PS points from each of the eight teams was validated.

1. A project was established in ArcGIS 9.1.
    a. The coordinate system of the data frame was set to the Lambert III cartographic reference system.

2. The orthorectified, 0.5m metre resolution, aerial photograph was loaded into the project to act as a reference for the locations of the PS points. This photograph was referenced to the Lambert III cartographic reference system.

3. Data from each of the eight processing teams was loaded into the project.
    a. This data had been pre-processed by BRGM. Pre-processing reprojected the data as well as applying a preliminary shift to the data. These pre-processing steps can be found in section T7.1.1b of the Annex.
    b. Pre-processed data was received from BRGM in a MapInfo format. This was converted to ESRI shapefiles using the Universal Translator function in MapInfo.

4. Locations suitable for the geocoding check were then found:
    a. All teams results were displayed over the orthophoto, each teams points were displayed in a different colour.
    b. Locations were then chosen based on points from each team being present and it being easy to identify what object the points were related to. The difficulty involved in relating PS points to an object, meant it was necessary to choose either features with a distinctive shape or features with a sufficient separation from other ground features, such as an electricity pylon or building set amongst trees.

5. Once a suitable location was found, only one team's data was displayed over the orthophoto. The user identified the points they thought related to the feature on the photograph.

6. A vector was then digitised from the points observed location to where it was thought the actual location should be. As this is a subjective process, there may be some errors involved in this process. Coordinates for the start and end of the vector were recorded.

7. Differencing the coordinates gave a shift, in meters, for both the X and Y directions between the point's observed and estimated actual locations.

8. These shifts were recorded in an Excel document for each location and for each team. For each measurement, a screen grab was made of the orthophoto location with overlaid points and shift vectors. Numerical information of the start and end coordinates of the the shifts was also added to a Word document.

9. The vector magnitudes and azimuths were then calculated from the X and Y shifts for each point. Then an overall vector resultant magnitude and azimuth was calculated for each Team, by summing all the Easting X and Northing Y shifts. It is important to realsise that these vector resultant values are not the same thing as independently determining the mean values of the X and Y shifts, and azimuths.

10. Finally in order to facilitate comparison between the Teams, standardised vector resultant shifts in the X and Y directions, and magnitudes, were calculated by dividing by the number of observations

Results

Measurements were made for several types of scatterer in several different land use settings. These included buildings in a semi-industrial area at (854540E, 134224N). A coal feed for a power station at (853856E, 134718N). Isolated buildings in a rural area at (866432E,124899N). A single house in a rural area at (853618E, 128956N). A farm house in another rural area at (852731E, 131849N). An electricity pylon on the edge of an agricultural field at (854462E, 131781N). Another electricity pylon beside a road at (855295E, 133557N). A large rural building surrounded by woodland at (853653E, 134087N). Another rural building at (849832E, 134464N). Two isolated large office or housing blocks near (846475E, 134871N). A large commercial building surrounded by trees and grass near (857253E, 131077N), and another large commercial building partially surrounded by houses at (847762E, 131265N).

Summary of Measured Shifts

Team 1 has been excluded due to the difficulty of estimating what feature a PS point is related to. Shifts for each individual Team are shown in the following tables and graphs

Point	Observed X	Observed Y	Actual X	Actual Y	X Shift (m)	Y Shift (m)	Magnitude (m)	Azimuth
1	853846.20	134710.46	853856.06	134718.15	-9.85	-7.69	12.50	232.01
1	853607.60	128956.54	853617.81	128957.73	-10.22	-1.19	10.29	263.35
1	852726.25	131844.96	852730.58	131848.54	-4.33	-3.58	5.62	230.45
1	854444.44	131782.75	854462.38	131780.65	-17.94	2.10	18.06	276.68
1	855286.44	133556.10	855296.57	133558.47	-10.13	-2.36	10.40	256.87
1	846402.18	134768.02	846397.47	134770.01	4.71	-1.98	5.11	112.87
2	846479.41	134868.52	846474.73	134871.30	4.68	-2.79	5.45	120.74
3	846324.28	134924.02	846318.86	134928.67	5.42	-4.66	7.15	130.66
1	847758.96	131262.99	847763.20	131262.87	-4.24	0.12	4.24	271.62
sum					-41.91	-22.03	78.82	
mean					-4.66	-2.45	8.76	
Std Dev					8.22	2.80	4.51	

Table 4.2.2 Summary of Team 2 measured shifts

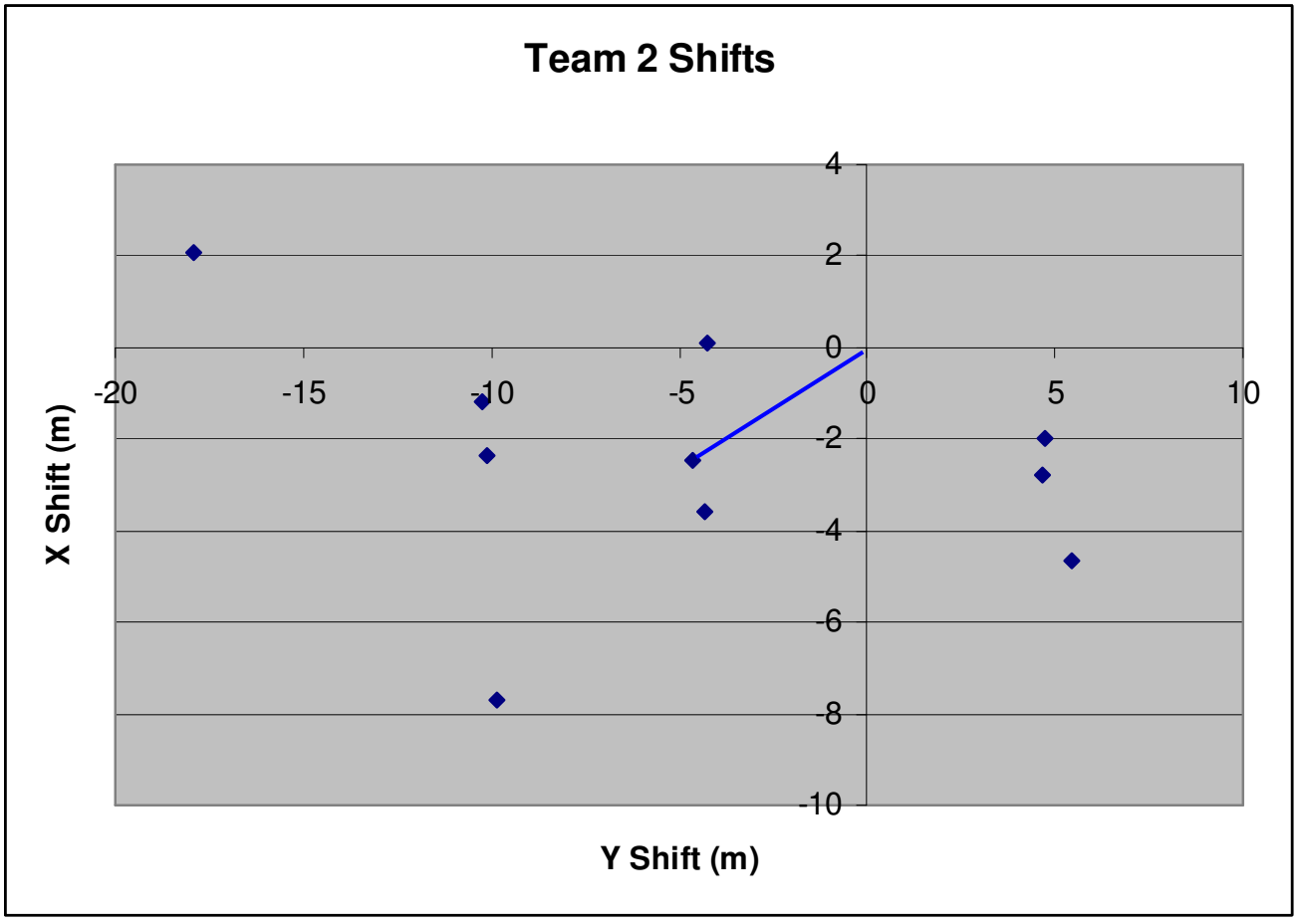


Figure 4.2.2 Graphic display of Team 2 shifts (blue line shows mean shift vector)

Table 4.2.3 Summary of Team 3 measured shifts

Point	Observed X	Observed Y	Actual X	Actual Y	X Shift (M)	Y Shift (M)	Magnitude	Azimuth
1	853615.84	128953.52	853617.76	128955.69	-1.91	-2.17	2.89	221.42
1	846327.86	134928.67	846322.08	134930.16	5.78	-1.49	5.97	104.41
1	847757.71	131265.14	847761.68	131265.29	-3.98	-0.15	3.98	267.87
sum					-0.11	-3.80	12.84	
mean					-0.04	-1.27	3.21	
Std Dev					5.14	1.03	1.56	

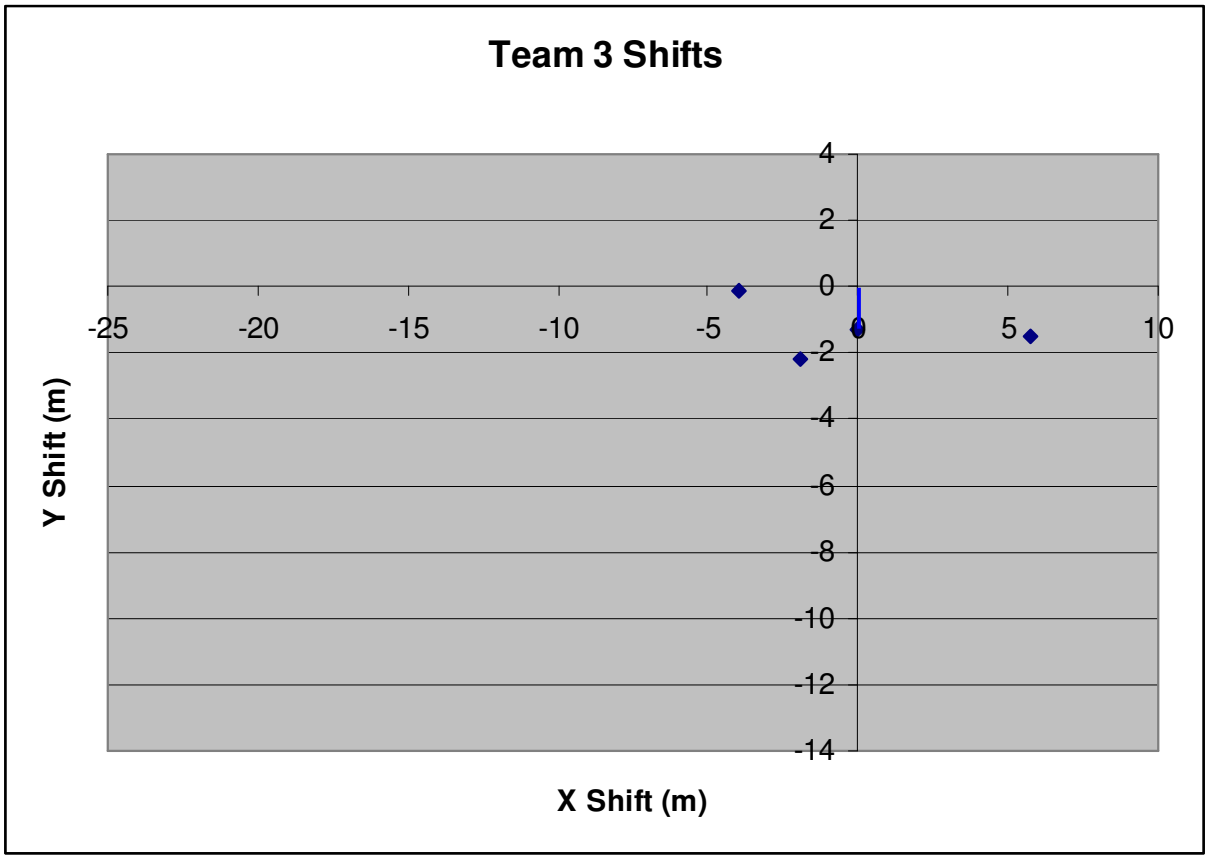


Figure 4.2.3 Graphic display of Team 3 shifts **(blue line shows mean shift vector)**

Table 4.2.4 Summary of Team 4 measured shifts

Point	Observed X	Observed Y	Actual X	Actual Y	X Shift (M)	Y Shift (M)	Magnitude	Azimuth
1	853856.39	134710.17	853855.09	134716.51	1.30	-6.34	6.47	168.41
1	866433.28	124896.05	866431.52	124899.25	1.76	-3.20	3.65	151.21
2	866492.45	124901.72	866490.99	124904.48	1.46	-2.76	3.13	152.11
3	866554.96	124915.08	866552.68	124919.19	2.28	-4.11	4.70	150.90
4	866554.12	124919.26	866552.09	124923.05	2.02	-3.79	4.30	151.90
1	853856.45	134712.40	853855.08	134716.32	1.37	-3.92	4.15	160.75
1	853616.04	128953.70	853617.04	128956.45	-1.00	-2.75	2.93	199.95
1	855290.08	133553.88	855294.68	133557.19	-4.60	-3.31	5.66	234.23
1	855910.07	134479.63	855909.18	134485.48	0.89	-5.85	5.92	171.31
1	846326.69	134924.74	846322.82	134928.03	3.88	-3.29	5.09	130.32
1	847758.50	131265.38	847761.70	131265.82	-3.19	-0.44	3.22	
sum					6.18	-39.76	49.21	
mean					0.56	-3.61	4.47	
Std Dev					2.51	1.57	1.20	

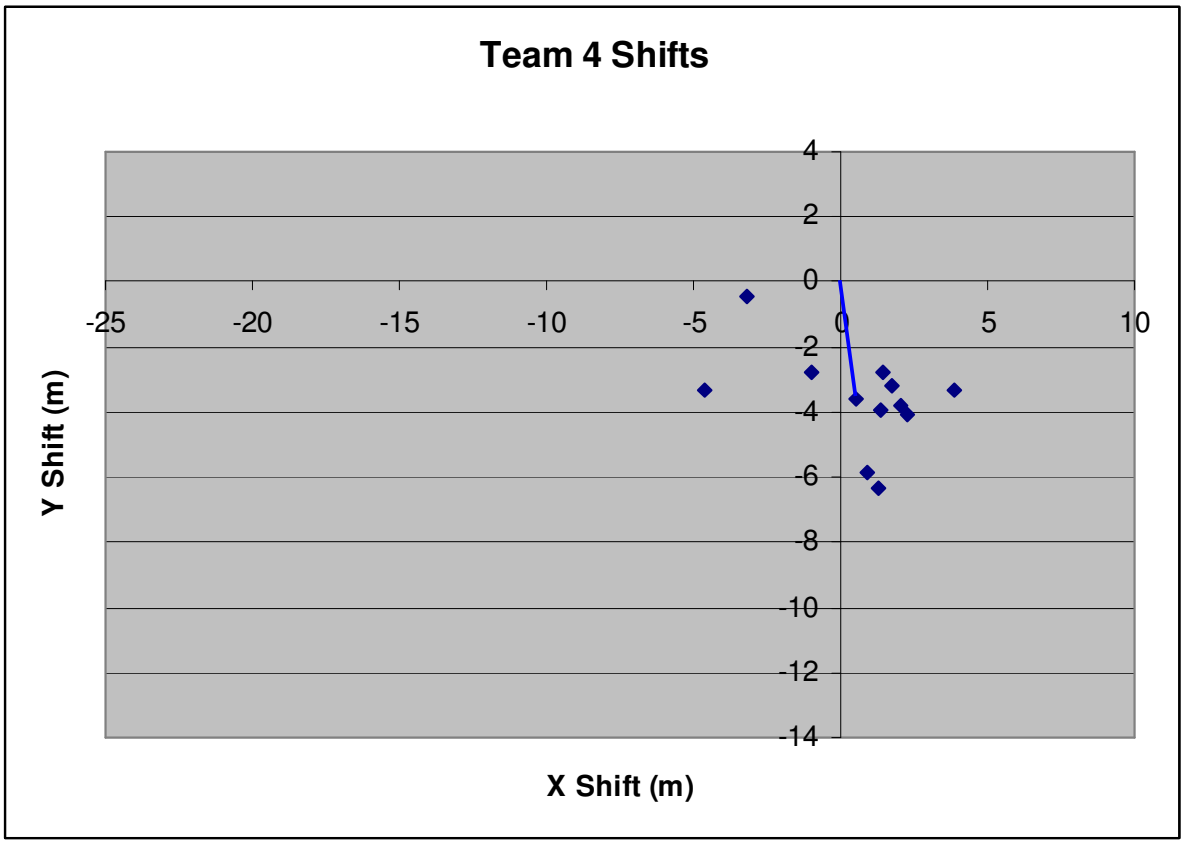


Figure 4.2.4 Graphic display of Team 4 shifts (blue line shows mean shift vector)

Table 4.2.5 Summary of Team 5 measured shifts

Point	Observed X	Observed Y	Actual X	Actual Y	X Shift (M)	Y Shift (M)	Magnitude	Azimuth
1	866552.88	124913.46	866551.40	124920.18	1.48	-6.72	6.88	167.57
1	846402.19	134777.55	846397.29	134782.45	4.89	-4.90	6.93	135.05
2	846477.66	134872.01	846471.12	134874.85	6.54	-2.84	7.13	113.44
1	847261.74	131075.14	847253.32	131076.97	8.42	-1.83	8.62	102.26
sum					21.34	-16.28	29.55	
mean					5.33	-4.07	7.39	
Std Dev					2.95	2.18	0.83	

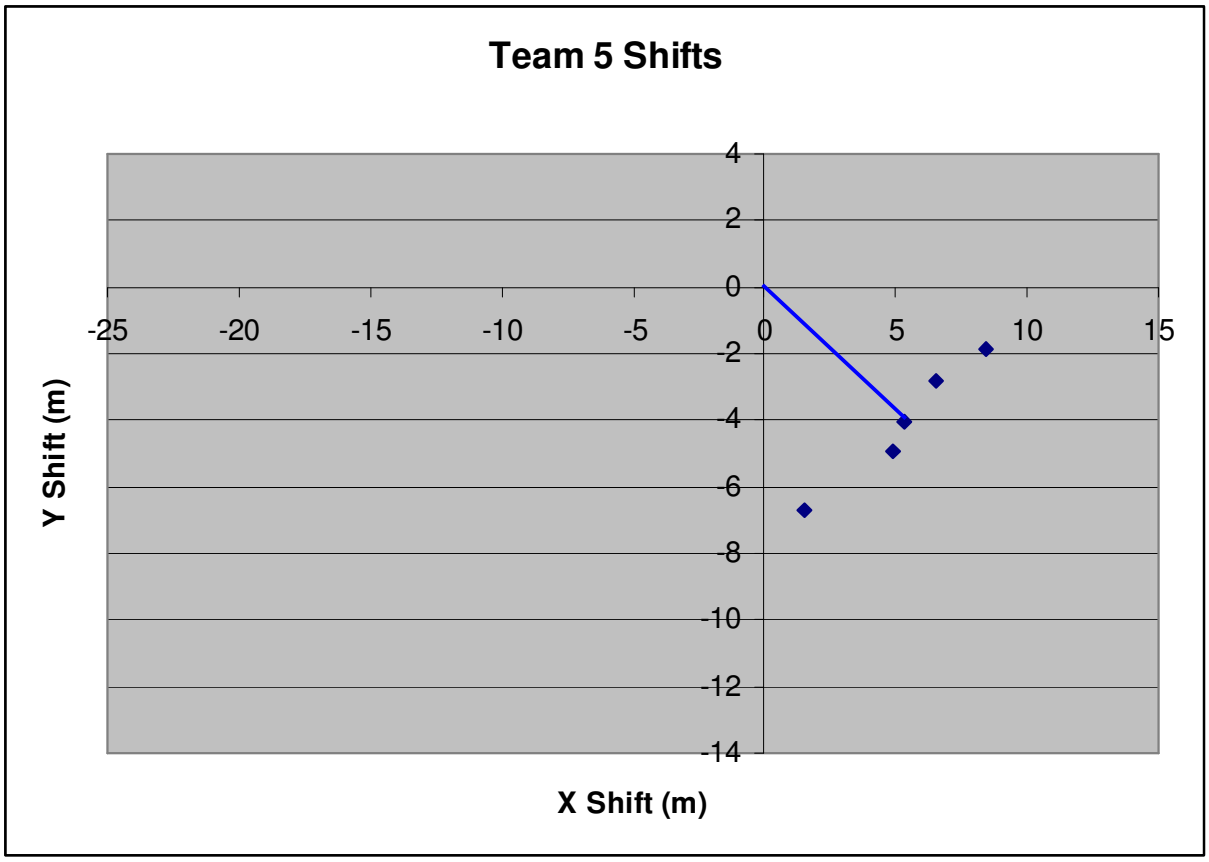


Figure 4.2.5 graphic display of Team 5 shift (blue line shows mean shift vector)

Table 4.2.6 Summary of Team 6 measured shifts

Point	Observed X	Observed Y	Actual X	Actual Y	X Shift (M)	Y Shift (M)	Magnitude	Azimuth
1	853828.07	134697.52	853847.77	134699.56	-19.71	-2.04	19.81	264.09
1	852717.71	131842.81	852736.81	131844.39	-19.10	-1.58	19.17	265.28
1	854442.20	131778.31	854461.84	131780.40	-19.65	-2.09	19.76	263.92
1	855282.37	133552.77	855295.76	133558.22	-13.39	-5.44	14.45	247.88
1	855886.40	134484.81	855910.26	134486.24	-23.86	-1.43	23.90	266.56
1	849818.94	134465.33	849828.83	134468.49	-9.88	-3.16	10.38	252.27
1	846375.41	134766.66	846394.27	134789.49	-18.87	-22.83	29.62	219.57
2	846448.33	134864.64	846467.92	134873.88	-19.59	-9.24	21.66	244.75
3	846300.85	134928.12	846314.51	134933.82	-13.66	-5.69	14.80	247.37
1	847232.33	131073.85	847237.94	131075.54	-5.61	-1.69	5.86	253.23
1	847737.97	131261.46	847761.73	131266.89	-23.75	-5.43	24.37	257.13
1	854465.59	134230.07	854485.11	134237.19	-19.52	-7.12	20.78	249.96
2	854523.05	134221.56	854540.32	134223.98	-17.28	-2.42	17.44	262.02
sum					-223.86	-70.17	241.99	
mean					-17.22	-5.40	18.61	
Std Dev					5.26	5.79	6.21	

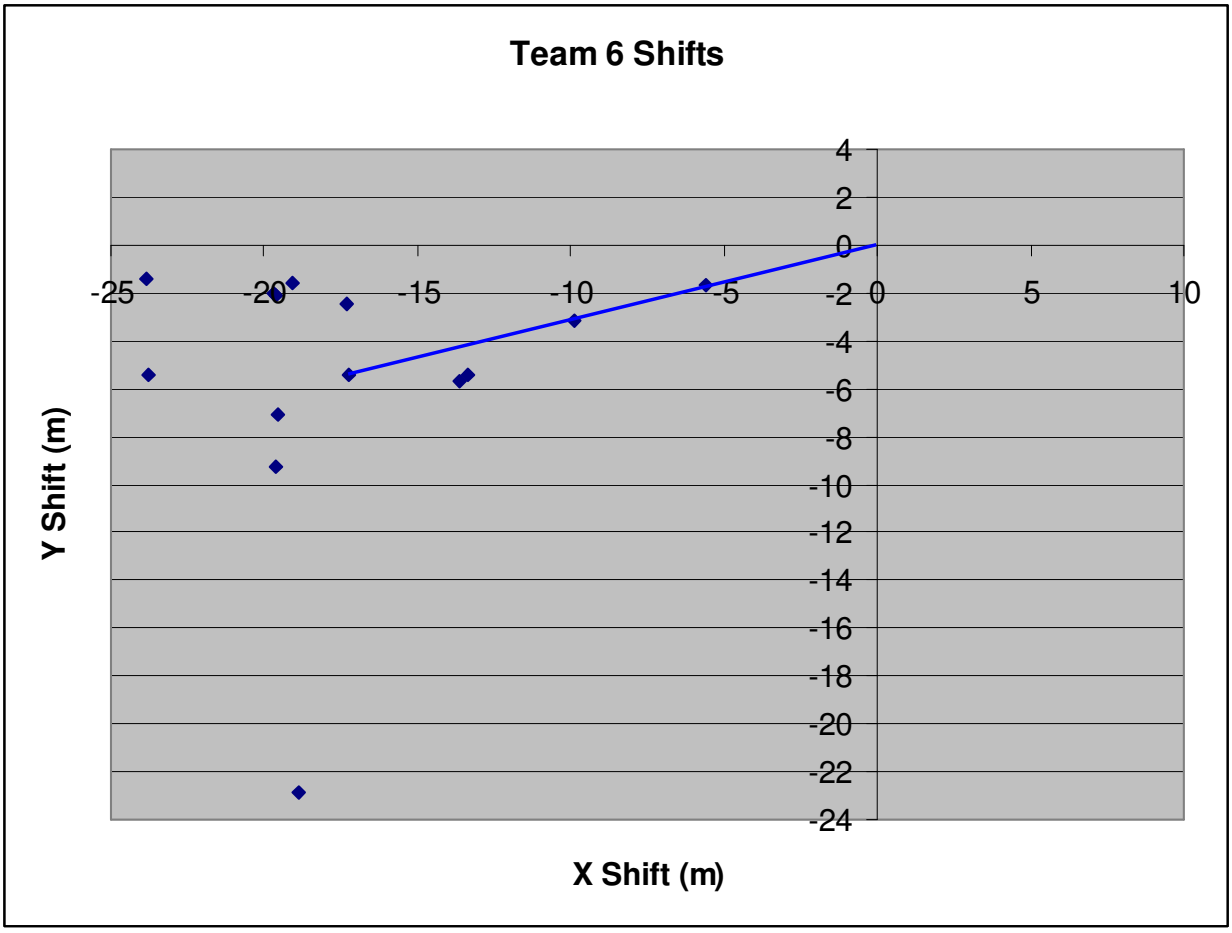


Figure 4.2.6 Graphic display of Team 6 shifts (blue line shows mean shift vector)

Table 4.2.7 Summary of Team 7 measured shifts

Point	Observed X	Observed Y	Actual X	Actual Y	X Shift (M)	Y Shift (M)	Magnitude	Azimuth
1	853858.41	134708.70	853854.72	134715.85	3.69	-7.15	8.04	152.71
1	866430.88	124897.29	866431.63	124904.12	-0.75	-6.83	6.87	186.30
2	866435.18	124892.16	866435.77	124899.92	-0.59	-7.76	7.78	184.35
3	866488.54	124898.61	866488.31	124903.81	0.22	-5.20	5.20	177.55
4	866491.31	124898.53	866490.74	124904.48	0.57	-5.95	5.98	174.52
5	866545.03	124917.39	866549.62	124922.45	-4.59	-5.06	6.83	222.18
6	866548.82	124920.62	866552.09	124924.79	-3.27	-4.17	5.30	218.11
7	866552.77	124915.88	866555.50	124920.80	-2.73	-4.93	5.63	209.01
8	866559.11	124910.69	866560.89	124914.52	-1.78	-3.83	4.22	204.91
9	866562.43	124910.22	866564.41	124914.99	-1.99	-4.77	5.16	202.62
1	855288.94	133551.00	855294.13	133556.41	-5.19	-5.42	7.50	223.76
1	855907.97	134480.22	855908.37	134484.97	-0.41	-4.75	4.77	184.87
1	853658.31	134086.59	853652.51	134087.28	5.80	-0.70	5.84	96.83
1	849836.32	134462.19	849831.61	134463.92	4.71	-1.72	5.02	110.09
1	846406.33	134774.52	846397.96	134778.20	8.37	-3.68	9.14	113.73
2	846481.07	134870.86	846471.67	134875.62	9.40	-4.76	10.54	116.85
3	846334.85	134922.92	846324.14	134927.99	10.71	-5.07	11.85	115.31
1	847259.45	131078.37	847253.42	131080.40	6.03	-2.03	6.37	108.61
1	847758.38	131260.88	847762.12	131262.10	-3.75	-1.21	3.94	252.04
1	854492.65	134232.22	854486.72	134237.94	5.93	-5.72	8.24	133.97
2	854548.61	134217.92	854541.39	134224.48	7.23	-6.56	9.76	132.22
sum					37.63	-97.25	143.98	
mean					1.79	-4.63	6.86	
Std Dev					4.94	1.91	2.14	

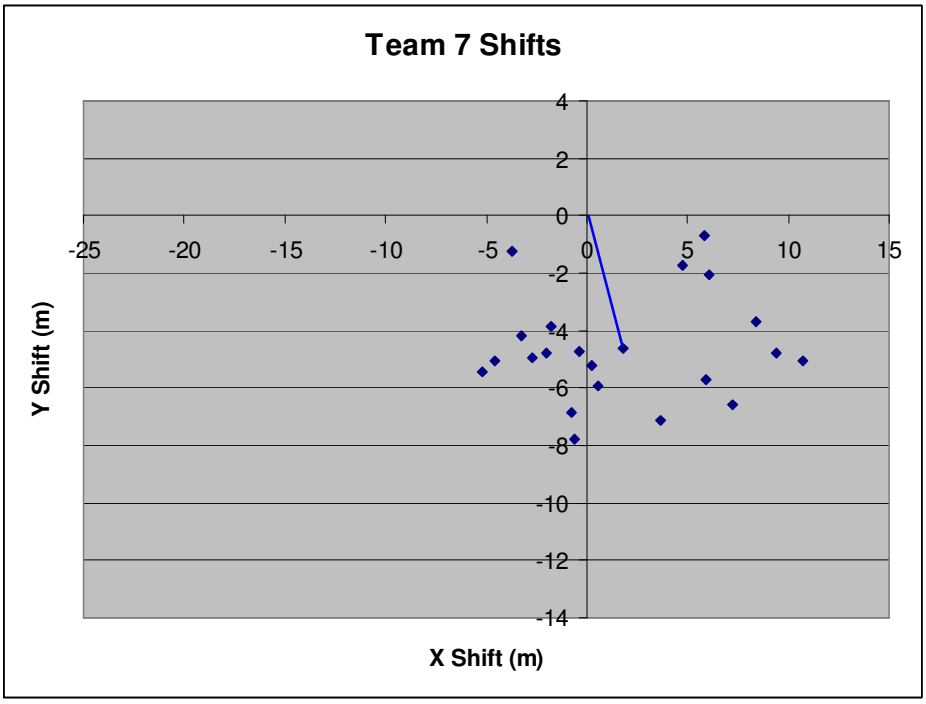


Figure 4.2.7 Graphic display of Team 7 shifts (blue line shows mean shift vector)

Table 4.2.8 Summary of Team 8 measured shifts

Point	Observed X	Observed Y	Actual X	Actual Y	X Shift (M)	Y Shift (M)	Magnitude	Azimuth
1	852701.97	131838.24	852737.11	131845.17	-35.14	-6.93	35.82	258.84
1	854466.18	131774.72	854461.59	131780.67	4.59	-5.95	7.52	142.34
1	855276.42	133538.65	855296.04	133558.48	-19.62	-19.83	27.90	224.70
sum					-50.17	-32.72	71.23	
mean					-16.72	-10.91	23.74	
Std Dev					20.02	7.74	14.60	

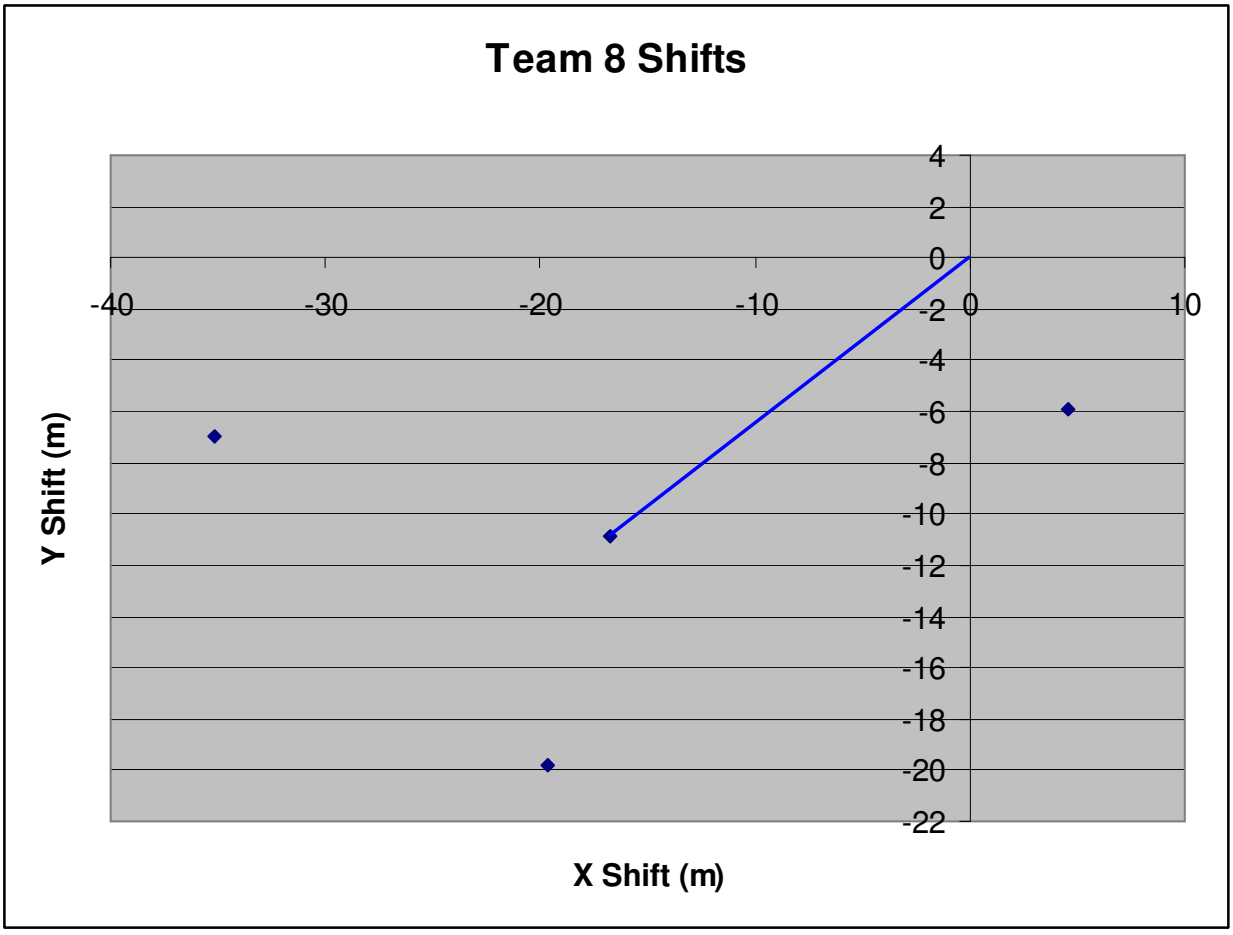


Figure 4.2.8 Graphic display of Team 8 shifts (blue line shows mean shift vector)

The measured shifts as given in the preceding tables can be combined and represented graphically as multiple vectors, shown in Figure3.1.15 below. The length of the line is the magnitude of the shift and the azimuth of the line is the bearing of the shift from north (the convergence point of all lines is the actual location of the PS point).

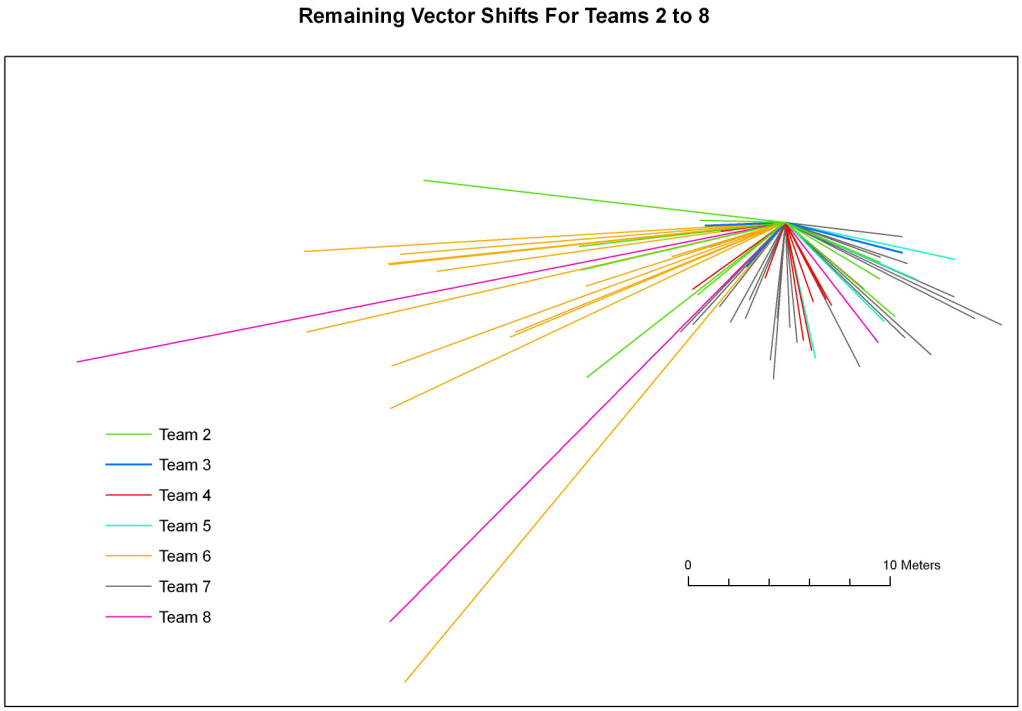


Figure 4.2.9 Summary plot showing the remaining vector shifts for Teams 2 to 8. Team 1 has been excluded due to the difficulty of estimating what feature a PS point is related to.

Comparison of results

The X (Easting) and Y (Northing) vector resultants for each team (excluding Team 1) have been summarised in Table 4.2.9 below. Also shown are the calculated magnitudes and directions (bearing from North) of the resultants. In order to compare results between the teams, standardised vector lengths were calculated and are plotted in Figure 4.2.10.

Table 4.2.9 Vector resultant shifts for each team in X and Y

	number points n	X vector resultant Xr (m)	Y vector resultant Yr (m)	Magnitude resultant R (m)	Resultant direction deg from N	Standardised X vector length Xr/n (m)	Standardised Y vector length Yr/n (m)	Standardised resultant length R/n (m)
T 2	9	-41.91	-22.03	47.35	242.27	-4.66	-2.45	5.26
T 3	3	-0.11	-3.80	3.80	181.66	-0.04	-1.27	1.27
T 4	11	6.18	-39.76	40.24	171.17	0.56	-3.61	3.66
T 5	4	21.34	-16.28	26.84	127.34	5.33	-4.07	6.71
T 6	13	-223.86	-70.17	234.60	252.60	-17.22	-5.40	18.05
T 7	21	37.64	-97.25	104.27	158.85	1.79	-4.63	4.97
T 8	3	-50.17	-32.72	59.90	236.89	-16.72	-10.91	19.97

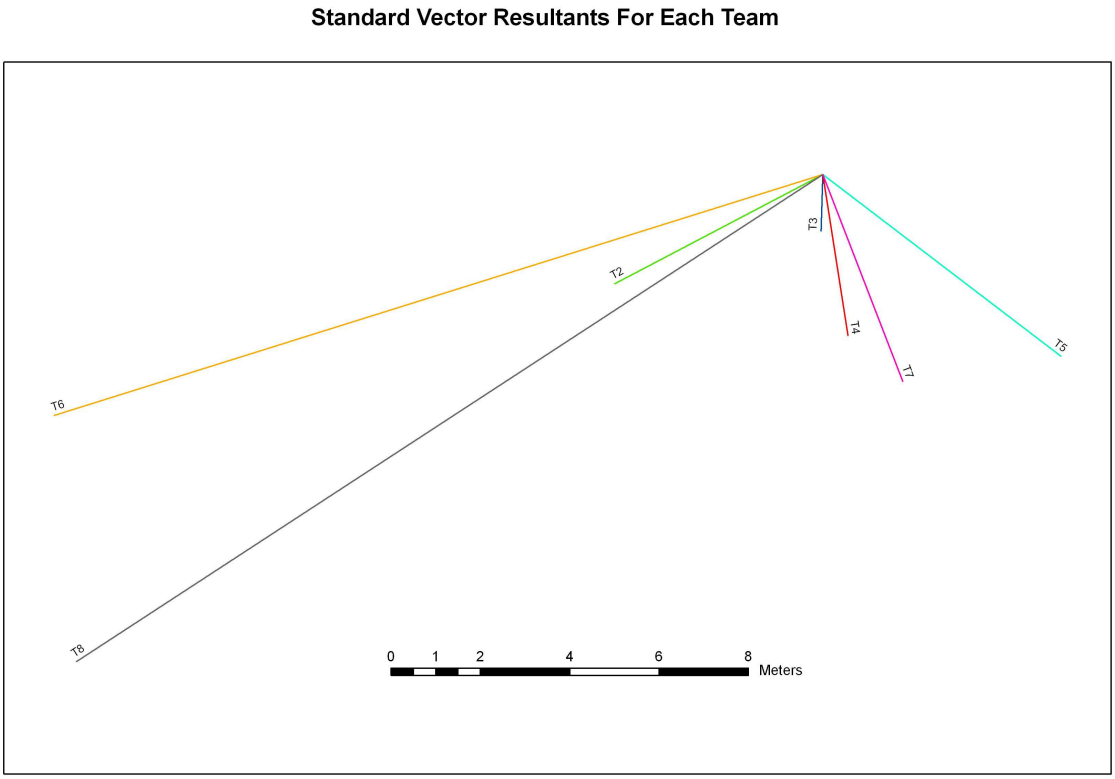


Figure 4.2.10 Graphic display of the standard vector resultants for Teams 2 to 8.

Conclusions for the check of geocoding quality

It was impossible to perform this check for Team One's data. The PS were presented in a gridded format making it very difficult to relate them to a ground feature.

Overall, the locational accuracy is good. However the reader must bear in mind that these measurements were made after a shift was applied to the data in the pre-processing steps, and relate to any remaining residual shifts.

In terms of the standardised vector resultant lengths, Team 3 has the best locational accuracy with a resultant length of 1.3m, with Teams 4, 7, 2, and 5 close behind, ranging from 3.7m to 6.7m. However these show a fairly wide distribution in the direction of the shifts, with Teams 7, 4 and 3 fairly tightly grouped with azimuths ranging from 159 to 182 degrees, while Team 5 is towards 127 degrees and Team 2 to 242 degrees.

The largest resultant vector shift is shown by Team 8 of 20m, on a bearing of 237 degrees. However this dataset is difficult to interpret as there are usually lots of PS points surrounding a feature, making it difficult to know which one relates to the feature. It should also be remembered that no shift was applied to Team 8 data during pre-processing. Team 6 showed the second largest shift of 18m towards 252 degrees, but interestingly, this was mostly due to a large westerly component.

It must be pointed out that there were many instances for many of the Teams where the PS points were thought to be in the correct location with probable zero displacement error.

Adding such zero X and Y shifts to the above individual and summary tables would inevitably reduce the mean shifts by an arbitrary amount. The conclusions above, thus relate to the likely maximum residual shifts after pre-processing adjustments.

The variation in shift directions for each individual teams results, for the various analysed points, indicates that the shifts are not the same from place to place, ie they are non-linear. These could be due to calculation of PS altitude errors, or perhaps even something to do with errors in the SAR orbit geometry. It is also noted that after the pre-processing corrections, there are few PS point shifts for any of the Teams to the North, which is confirmed by the summary of the standardised vector resultants in Figure 4.2.10.

4.3 Time series validation (TNO)

The purpose of the time series validation is to compare the PS-InSAR results on the cumulative displacement in time with the cumulative displacement in time as observed at levelling location and draw conclusions on how well these two datasets compare. Two approaches are followed:

1. a qualitative approach by plotting spatio-temporal profiles of the cumulative displacement along levelling sections
2. a statistical approach in which the average difference of the PS-InSAR time series and the levelling time series is quantified by means of a Root Mean Squared Error and in which the hypothesis is tested that the PS-InSAR time series and levelling time series are samples drawn from the same population

Spatio-temporal profiles

The PS-InSAR time series can spatially and temporally be compared with the levelling data along the the transects of the levelling lines. As levelling data the temporally interpolated data have been used so that at every InSAR-acquisition date we have an estimated levelling displacement. As PS-InSAR data the spatially interpolated data have been used so that at every leveling location we have a time series.

The results for the lines AXE and SNCF are shown in figures 4.3.1-18. These levelling sections have been chosen because the levelling of these lines started before the beginning of the InSAR-acquisition (15 july 1992), thus making it possible to adjust the time series (levelling and PS-InSAR) to a common zero reference and because these lines cover a sufficiently large stretch to visualize the subsidence. The figures show the levelling transect as discrete levelling points on the horizontal axis and time as acquisition date on the vertical axis. The cumulative displacement is colour coded.

AXE
This levelling line shows a distinctive subsidence process, starting in the first quarter of 1993 and ending in july 2002. Subsidence bowl starts to form at AXE 3-4050-1 and travels southward along the levelling section in conjunction with the progression of the mining works. Instantaneous subsidence velocities amount here to a maximum of about 12 cm/yr. The subsidence process can be followed along a stretch of about 4 km. Figures 4.3.2 – 9 show the results of the various processing chains. Apart from processing chain 8, none of the processing chains shows a full coverage of the section where the subsidence has taken place. Two of the processing chains (3 and 6) fail to capture any signal of subsidence along this levelling section. Outside the stretch where subsidence takes place, all processing chains show a fairly good to good coverage. Probably, part of the explanation for this has to do with the setting of quality thresholds which does not permit to accept PS-estimates for this stretch of levelling line because of the high subsidence rates for 7 of the 8 processing chains. All processing chains underestimate the displacement in time along the subsiding stretch. The underestimation of the subsidence is largest for that part of the subsidence starting before july 1995. Probably this has to do with the fact that there

is a large gap in SAR-acquisition between 1993 and 1995 thus posing problems with phase unwrapping for points showing large displacements. However, also the part of the section where subsidence starts well after july 1995, an underestimation of the displacement in time can be seen, though the onset of subsidence is visible.

SNCF

This levelling line shows two stretches of subsidence. One stretch about 3 km long and a smaller one of about 1 km, however not fully covered in time. Figures 4.3.11-18 show the results of the various processing chains. Again apart from processing chain 8, none of the processing chains shows a full coverage of the sections where the subsidence takes place and two of the processing chains (3 and 6) fail to capture any signal of subsidence along this levelling section. Processing chain 8 fully captures both areas where subsidence takes place, however the displacements measured by this processing chain are much less than what has been measured by conventional levelling. Again the stretches where subsidence starts well after the large gap in SAR-acquisition (1993-1995) is better covered and shows less underestimation of the displacement. Processing chains 1, 2, 5 and 7 capture signals of subsidence of both subsiding stretches but differ in the amount of final displacement with the levelling observations by more than 100 mm. Processing chain 4 only captures a couple of points of the large subsiding stretch.

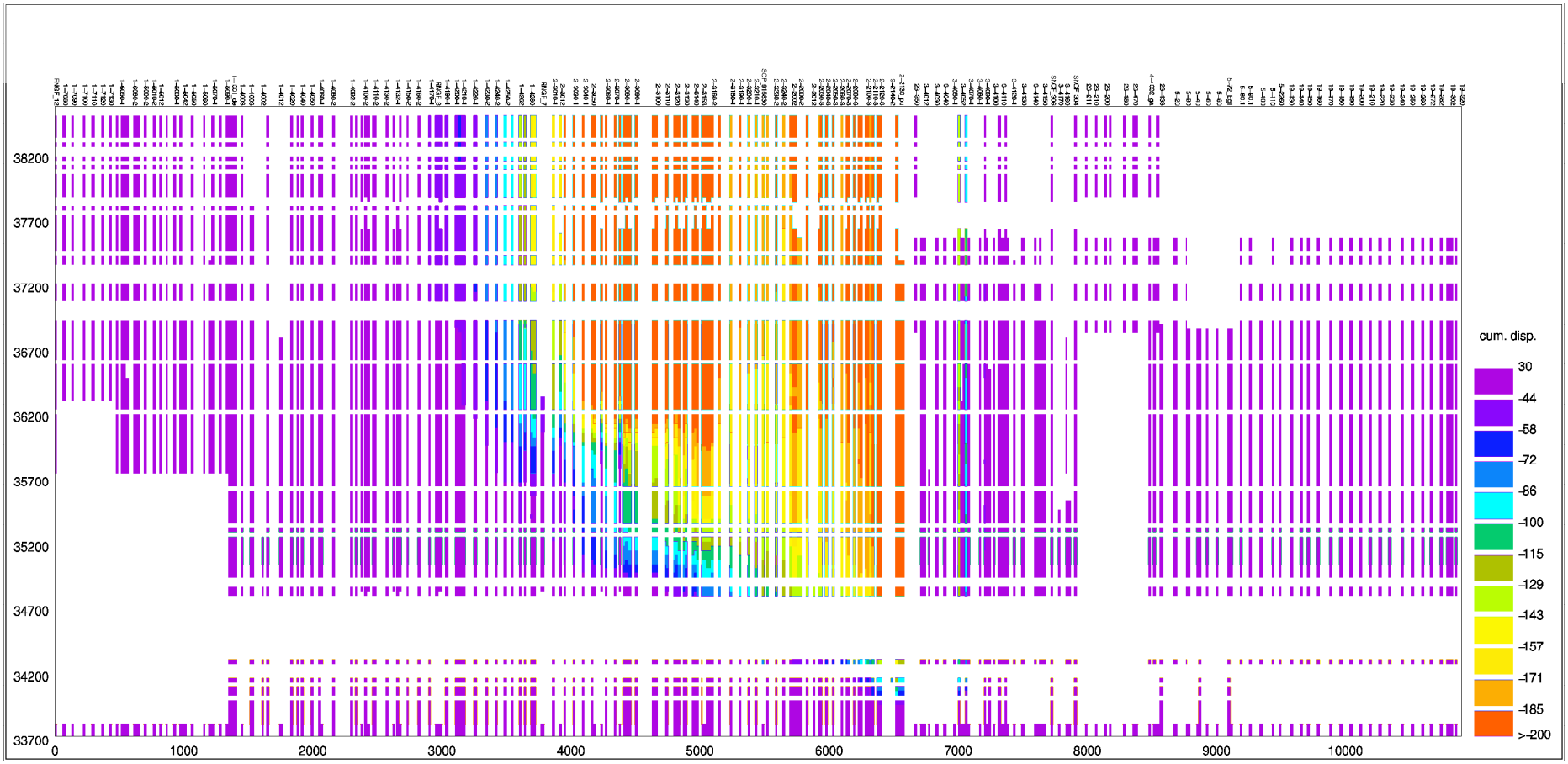


Figure 4.3.1 – Spatio-temporal profile of the levelling data along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

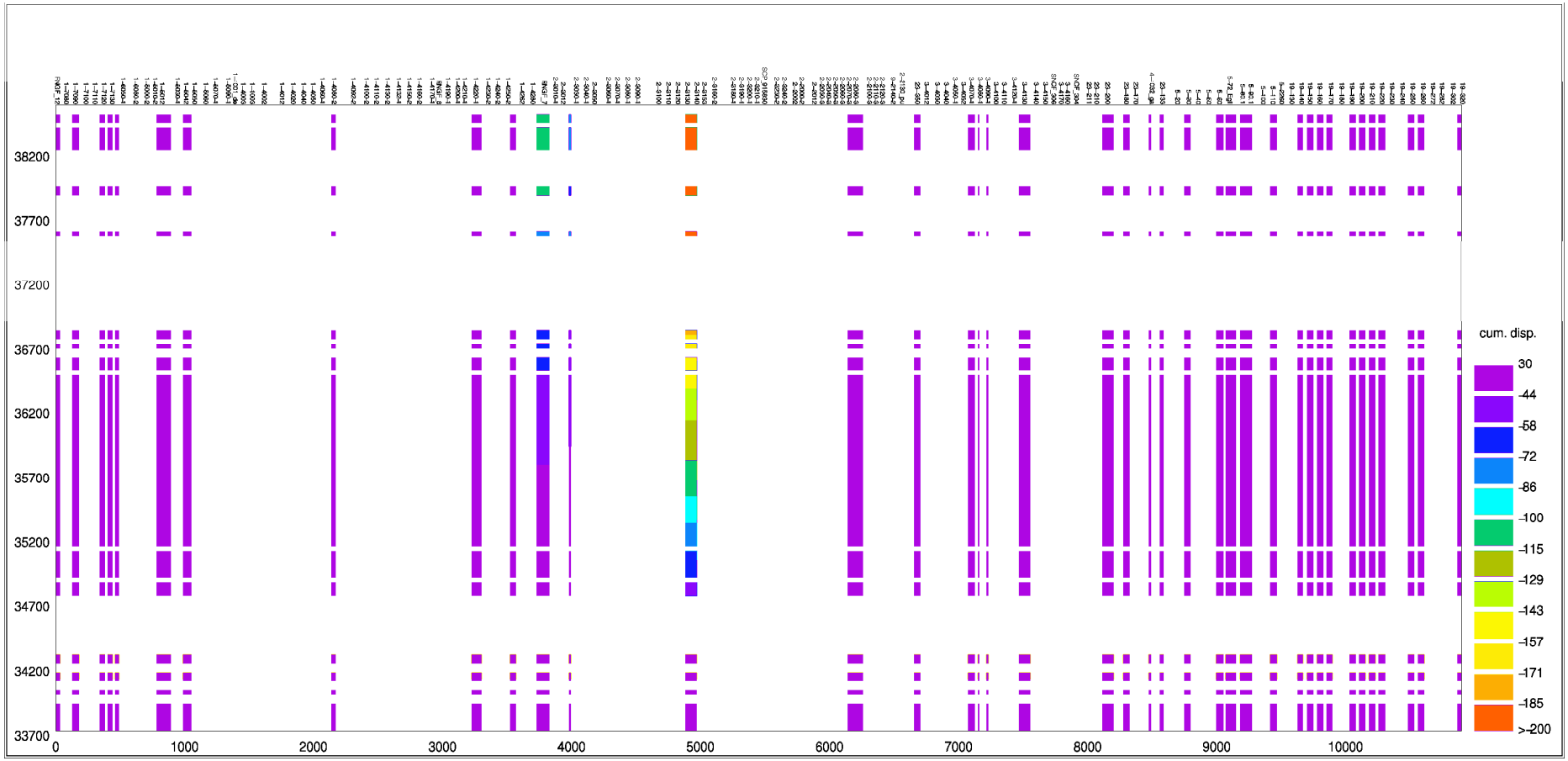


Figure 4.3.2 – Spatio-temporal profile of the PS-InSAR observations of processing chain 1, interpolated along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

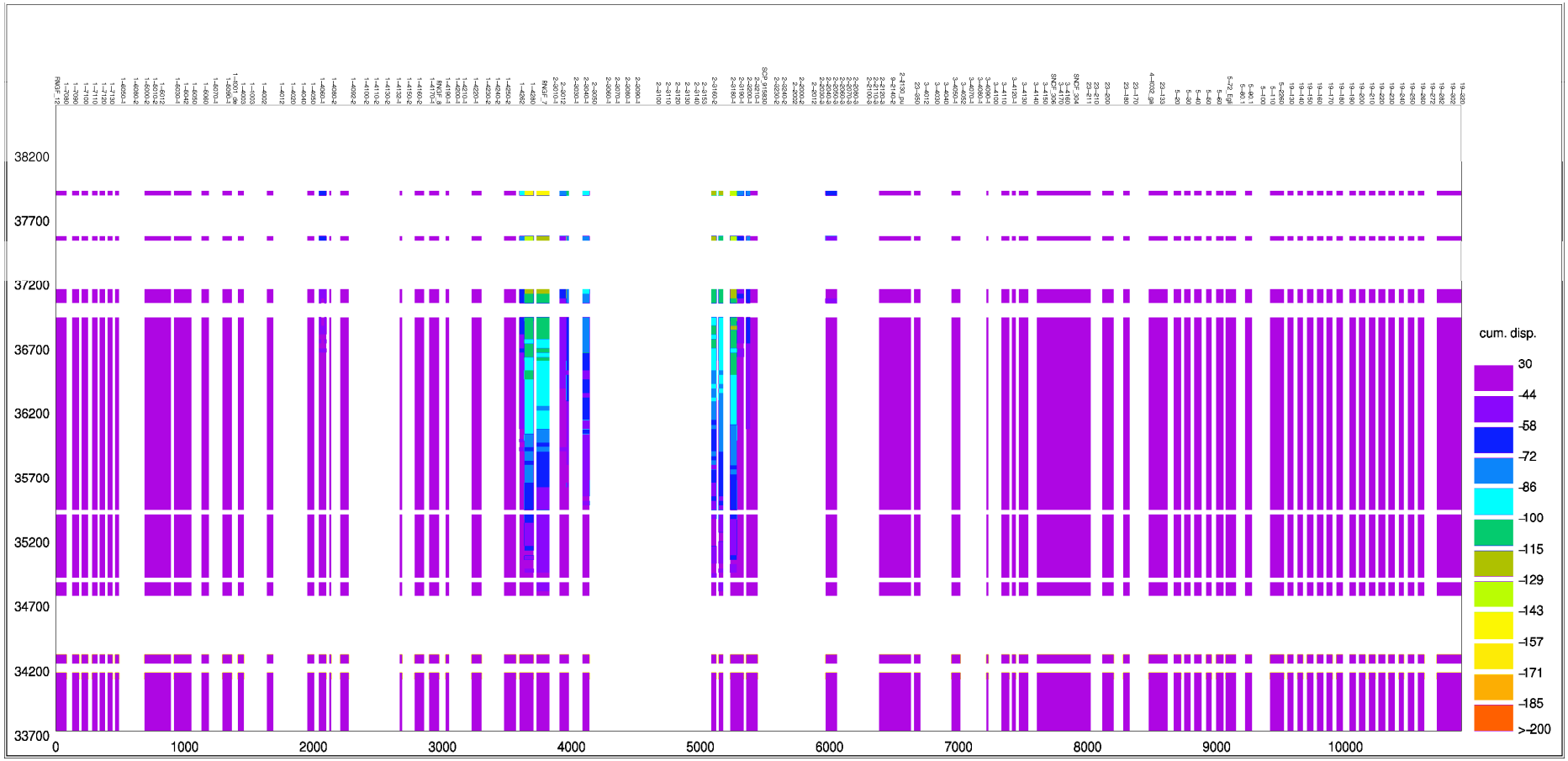


Figure 4.3.3 – Spatio-temporal profile of the PS-InSAR observations of processing chain 2, interpolated along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

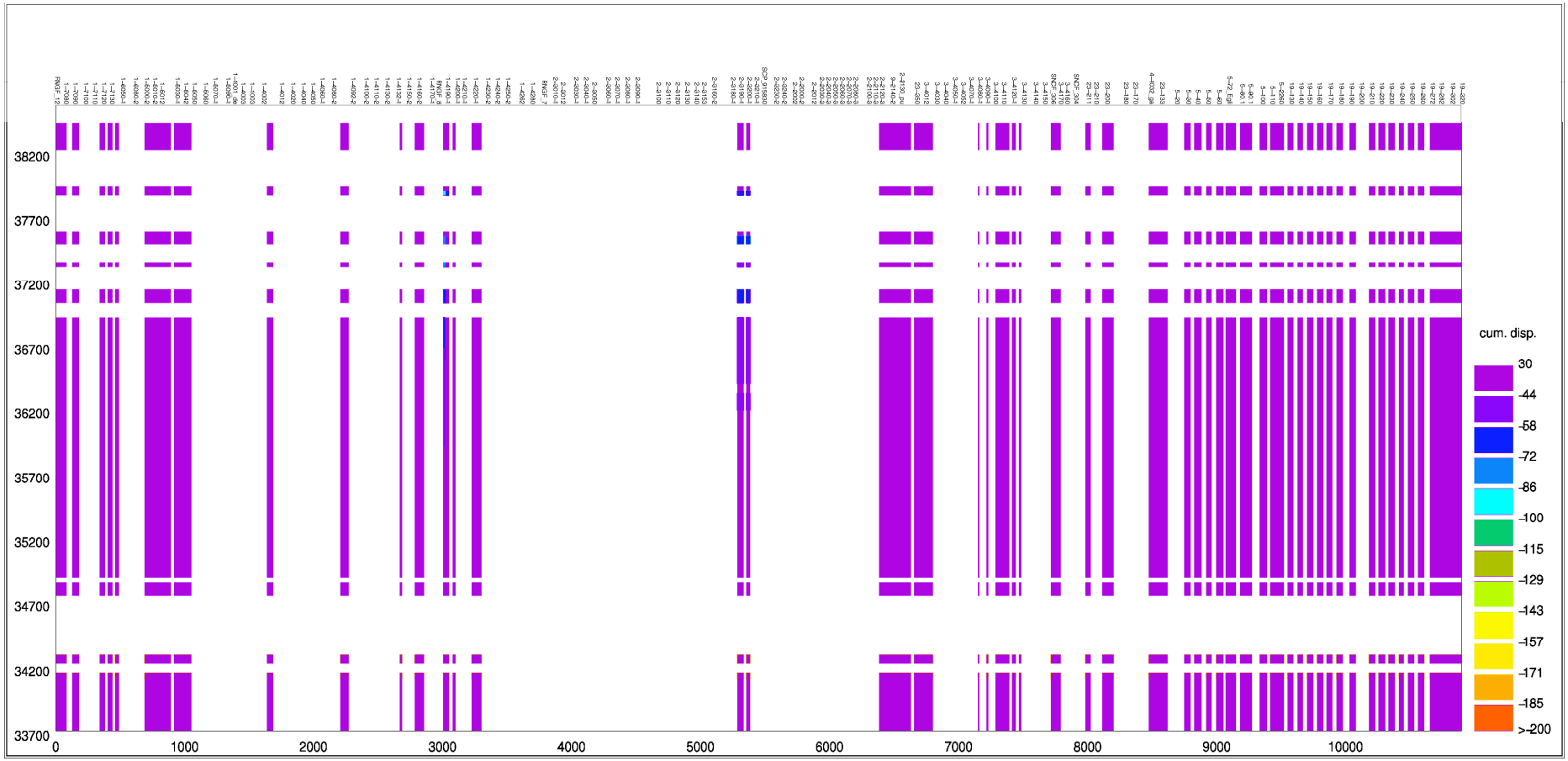


Figure 4.3.4 – Spatio-temporal profile of the PS-InSAR observations of processing chain 3, interpolated along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

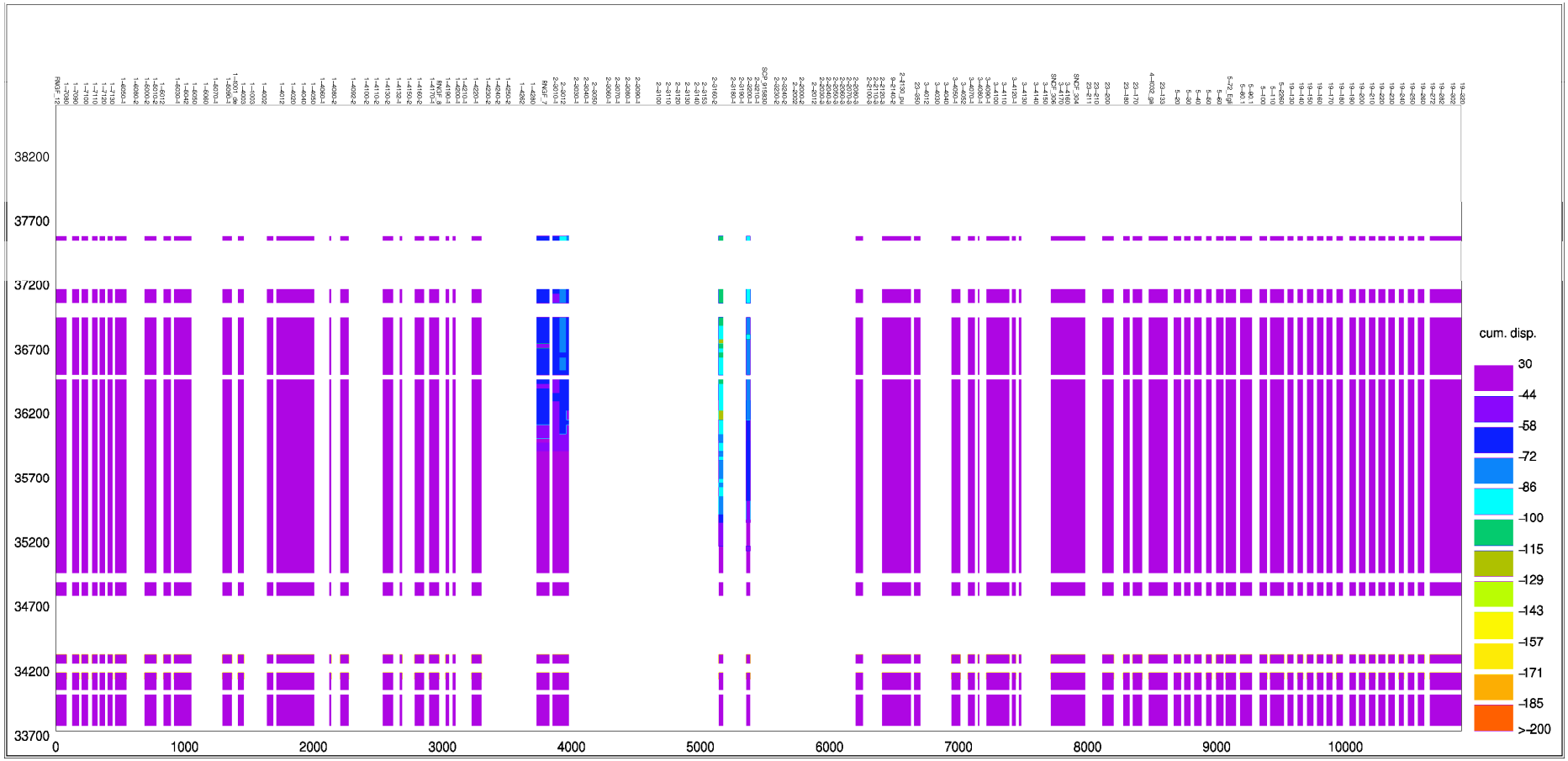


Figure 4.3.5 – Spatio-temporal profile of the PS-InSAR observations of processing chain 4, interpolated along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

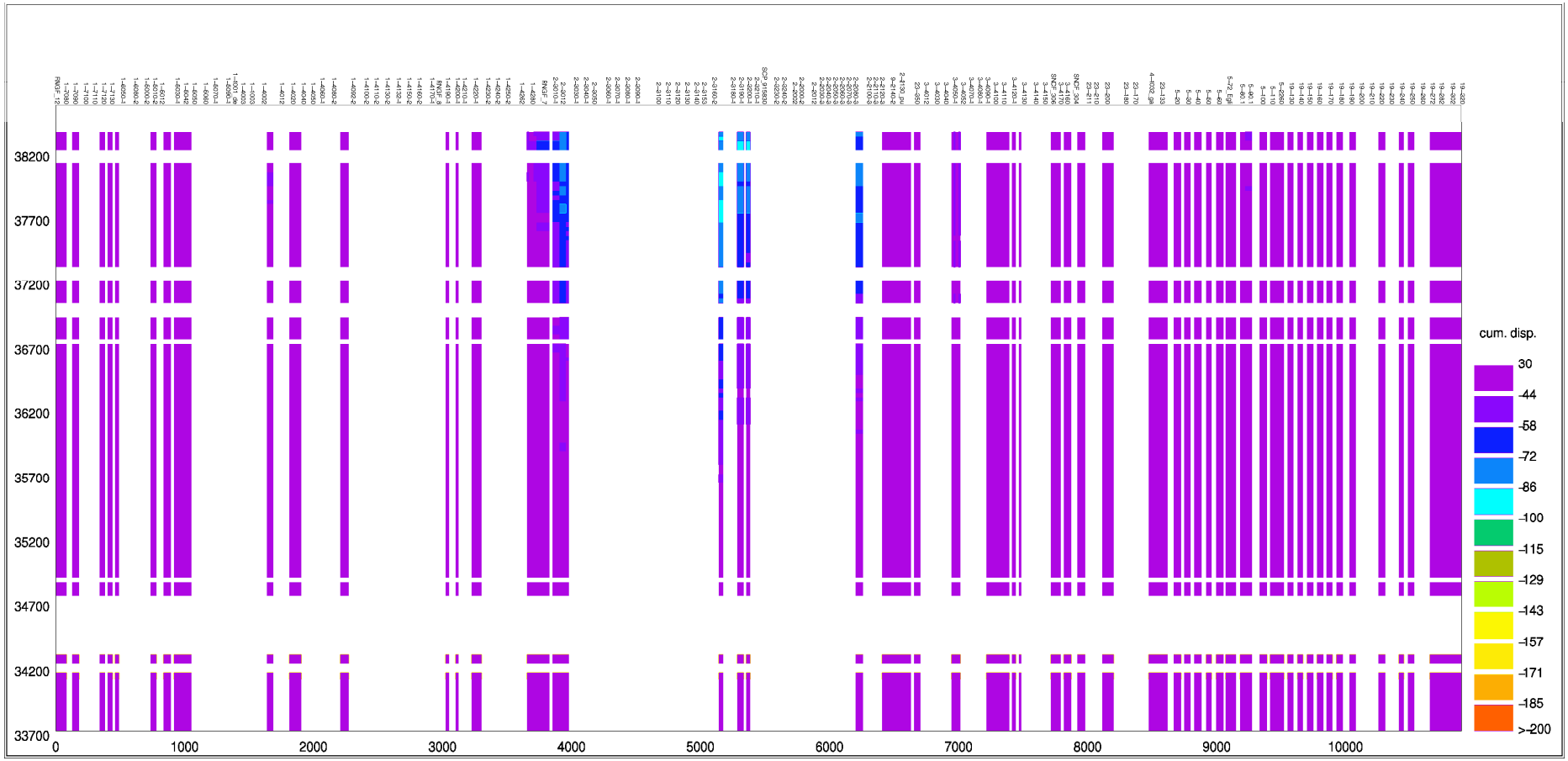


Figure 4.3.6 – Spatio-temporal profile of the PS-InSAR observations of processing chain 5, interpolated along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

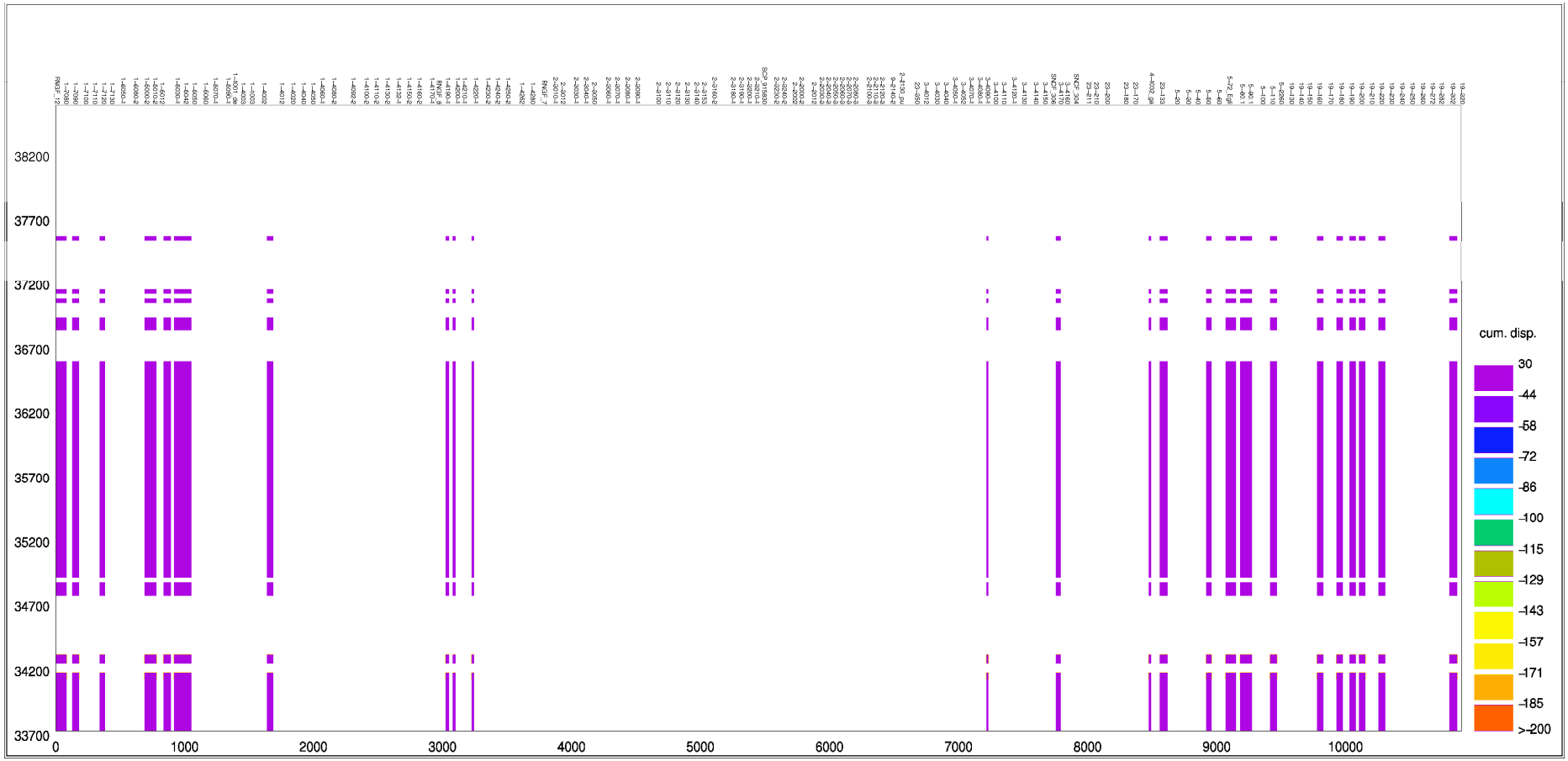


Figure 4.3.7 – Spatio-temporal profile of the PS-InSAR observations of processing chain 6, interpolated along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

Figure 4.3.8 – Spatio-temporal profile of the PS-InSAR observations of processing chain 7, interpolated along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

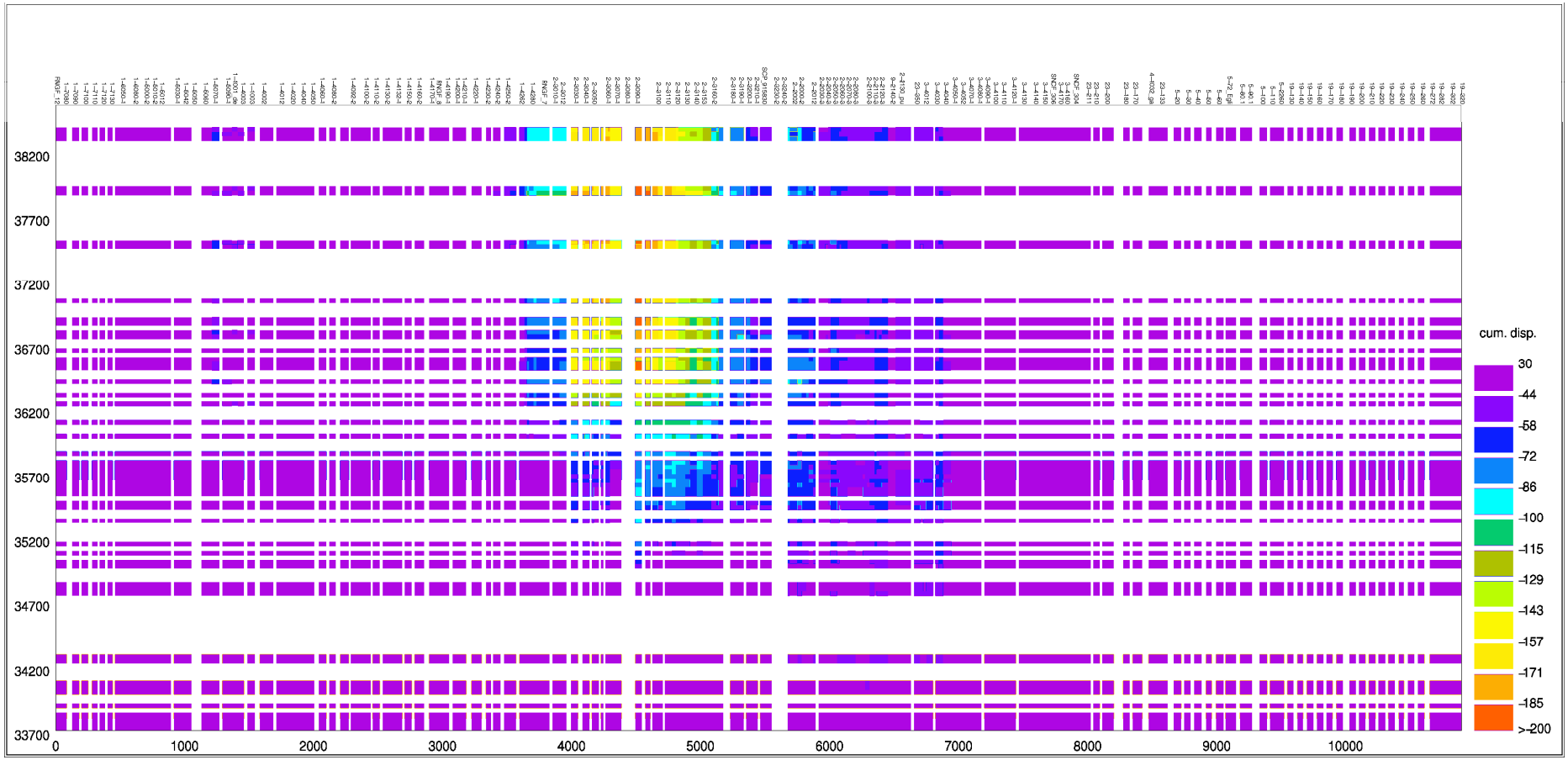


Figure 4.3.9 – Spatio-temporal profile of the PS-InSAR observations of processing chain 8, interpolated along levelling section AXE (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

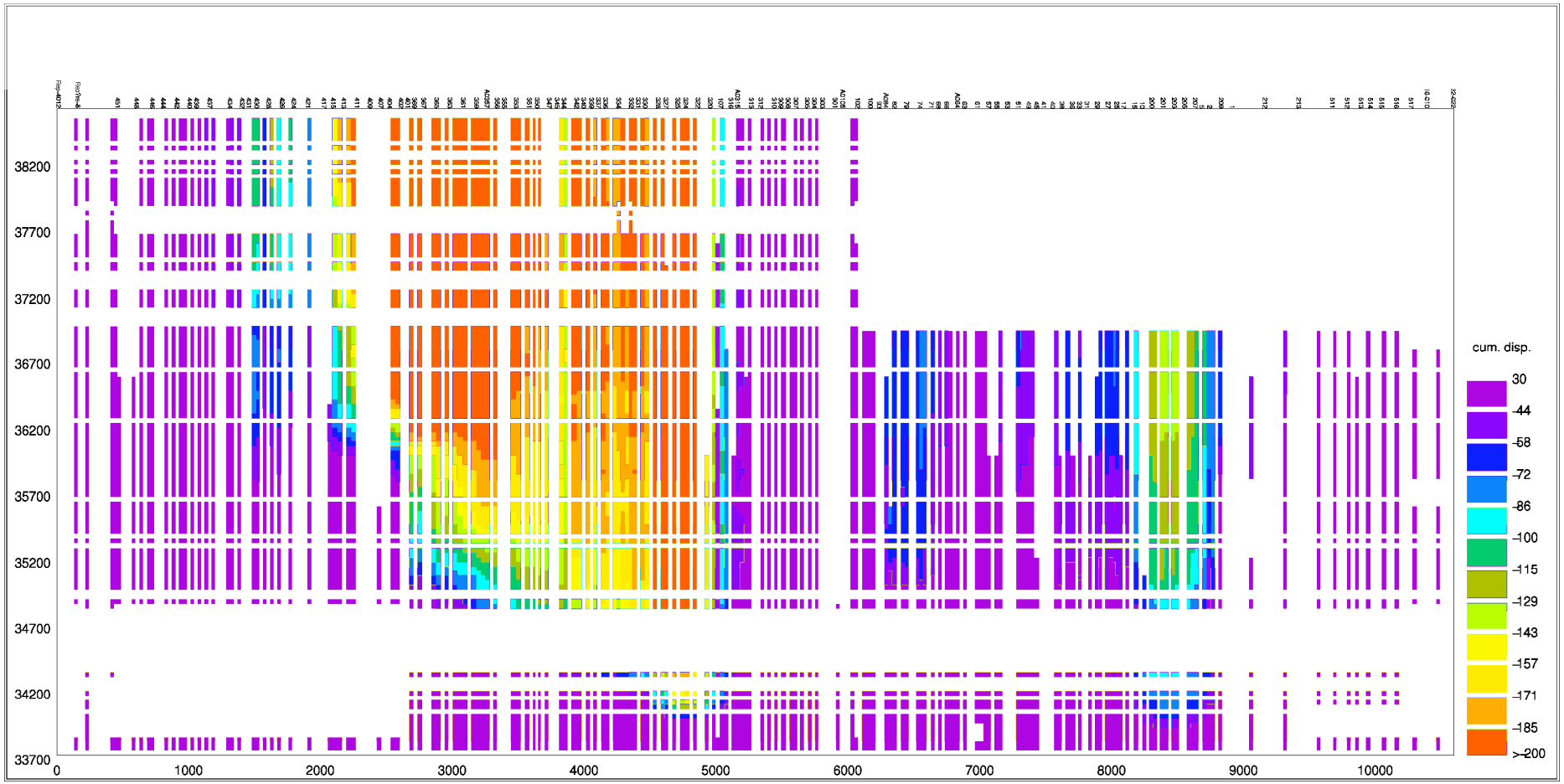


Figure 4.3.10 – Spatio-temporal profile of the levelling data along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

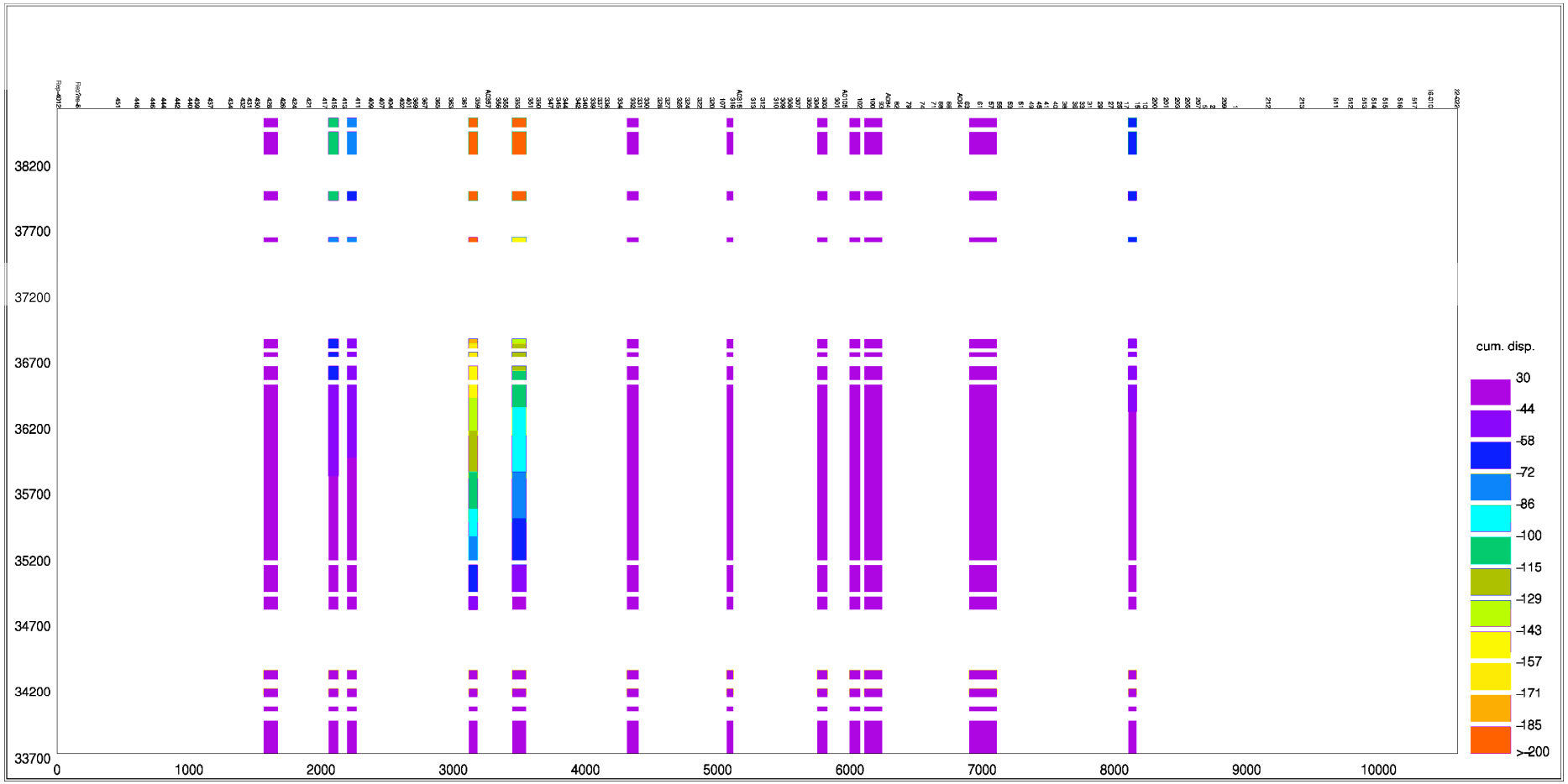


Figure 4.3.11 – Spatio-temporal profile of the PS-InSAR observations of processing chain 1, interpolated along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

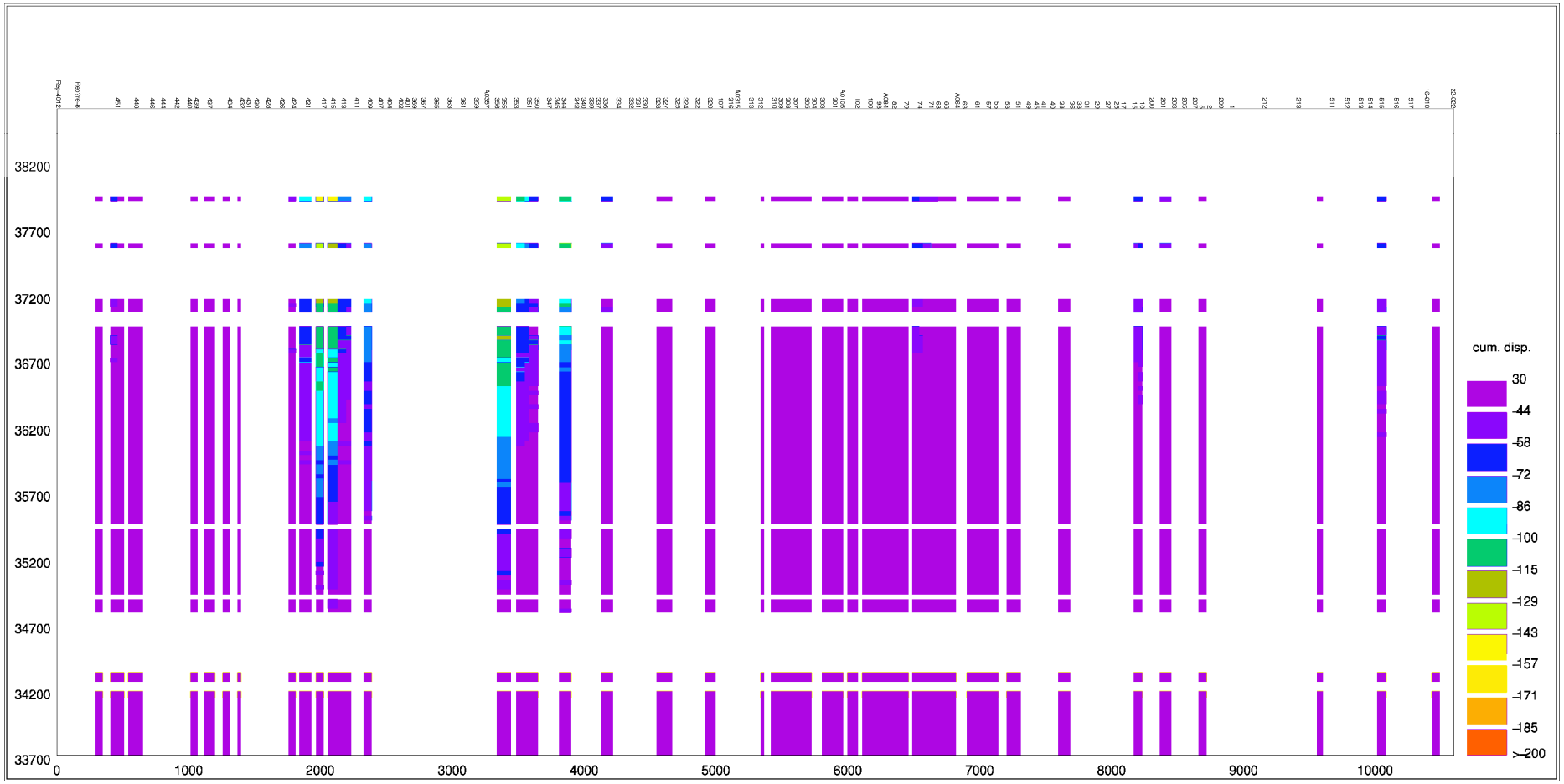


Figure 4.3.12 – Spatio-temporal profile of the PS-InSAR observations of processing chain 2, interpolated along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

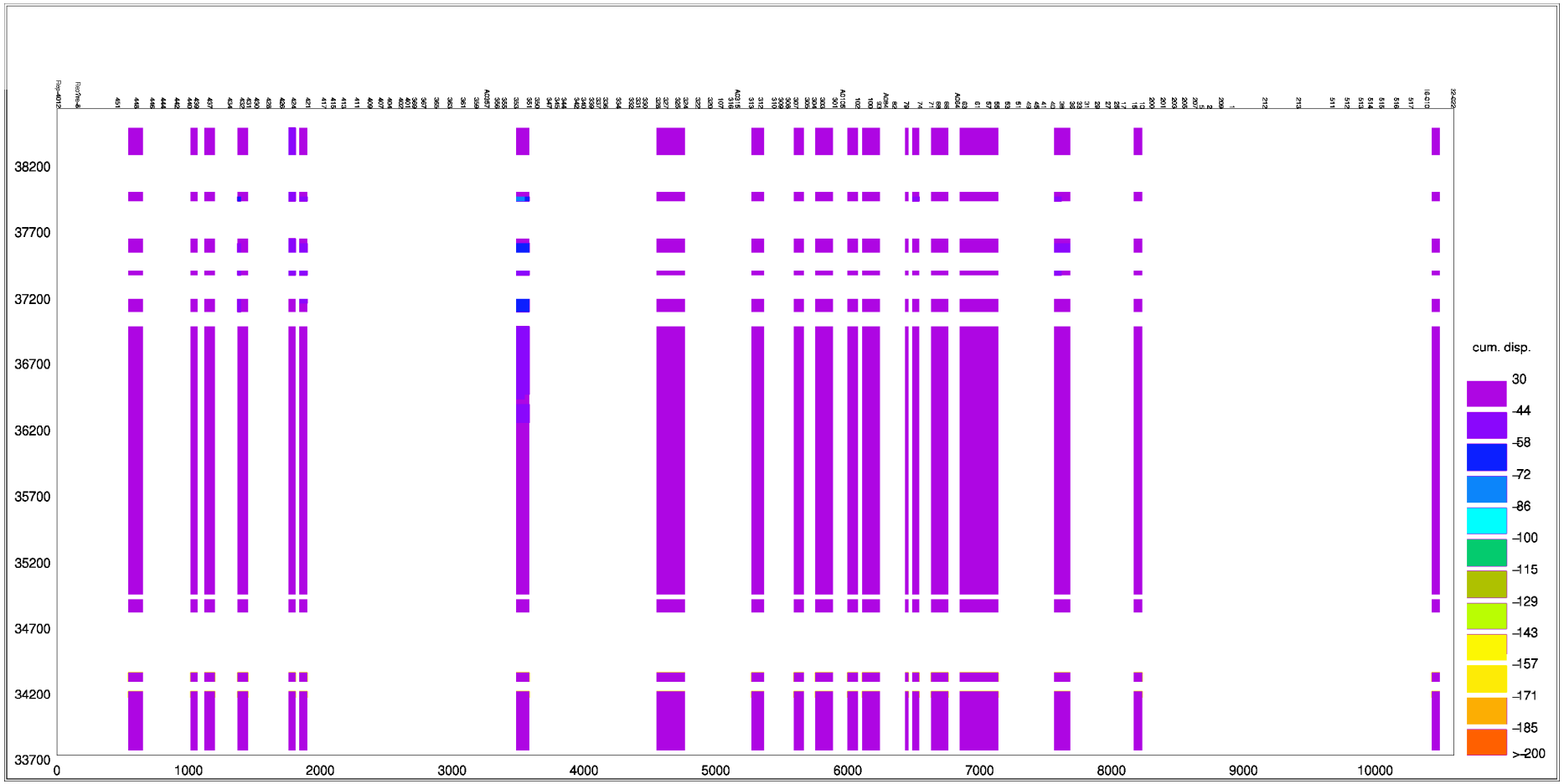


Figure 4.3.13 – Spatio-temporal profile of the PS-InSAR observations of processing chain 3, interpolated along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

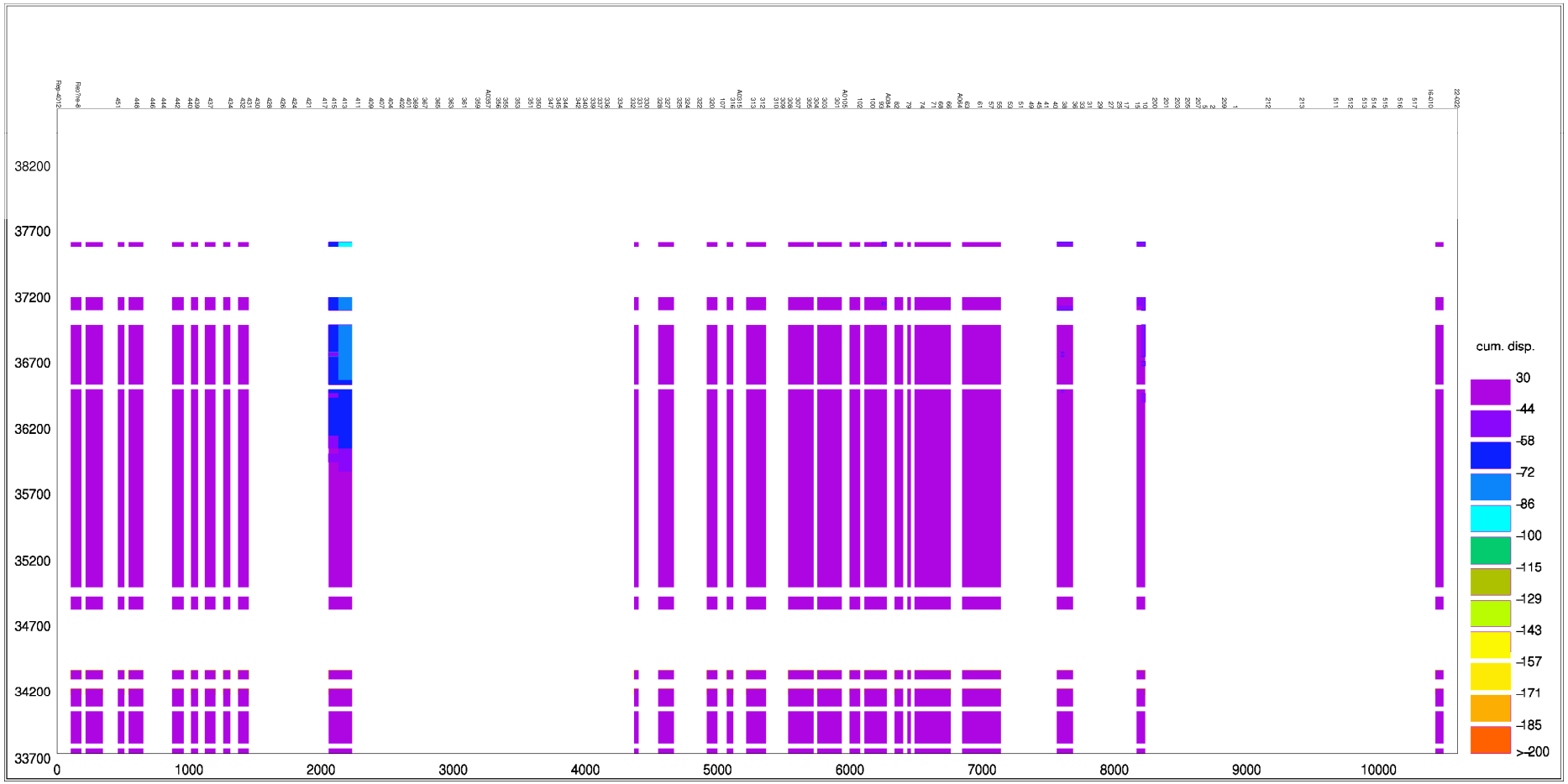


Figure 4.3.14 – Spatio-temporal profile of the PS-InSAR observations of processing chain 4, interpolated along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

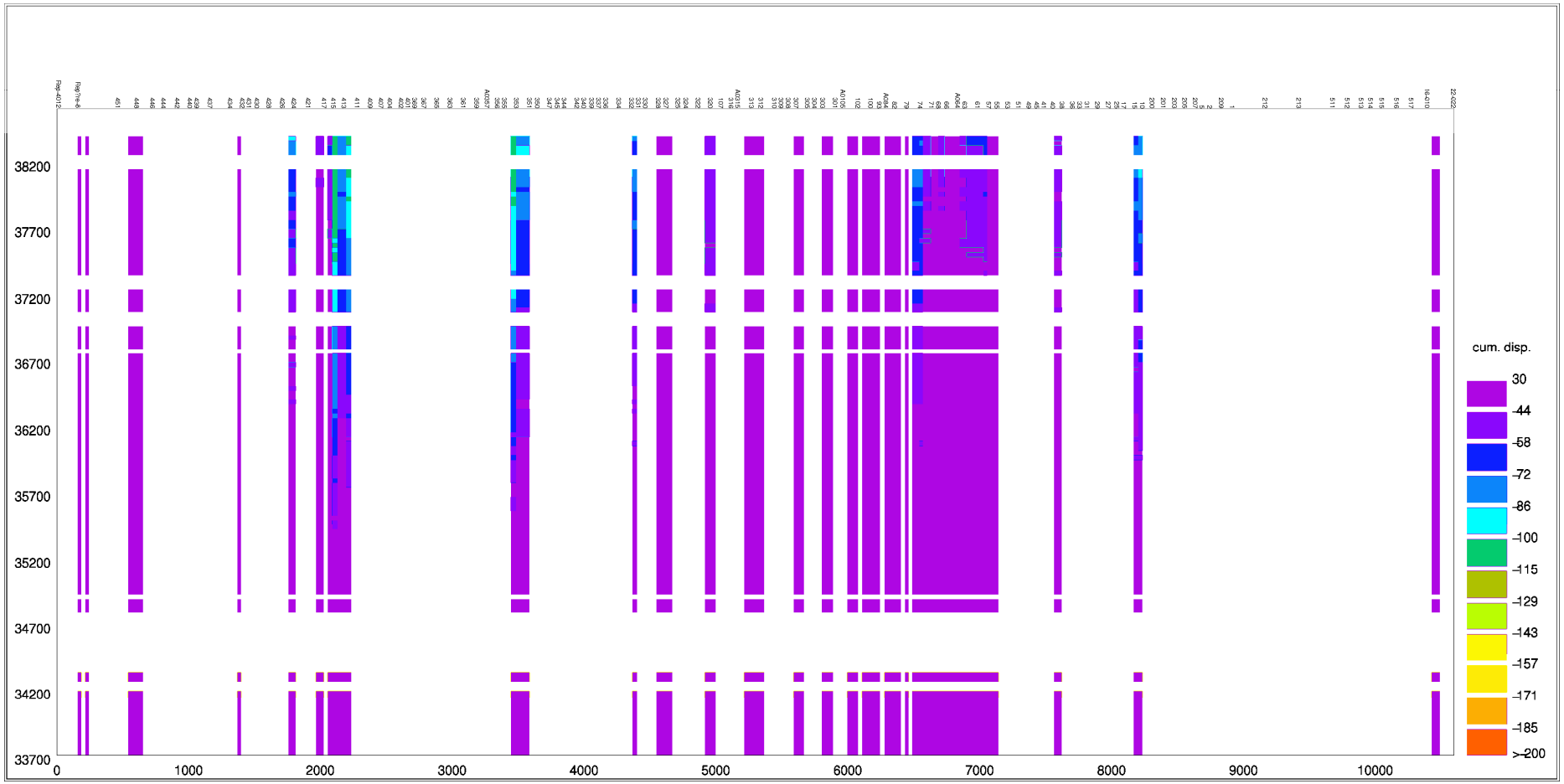


Figure 4.3.15 – Spatio-temporal profile of the PS-InSAR observations of processing chain 5, interpolated along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

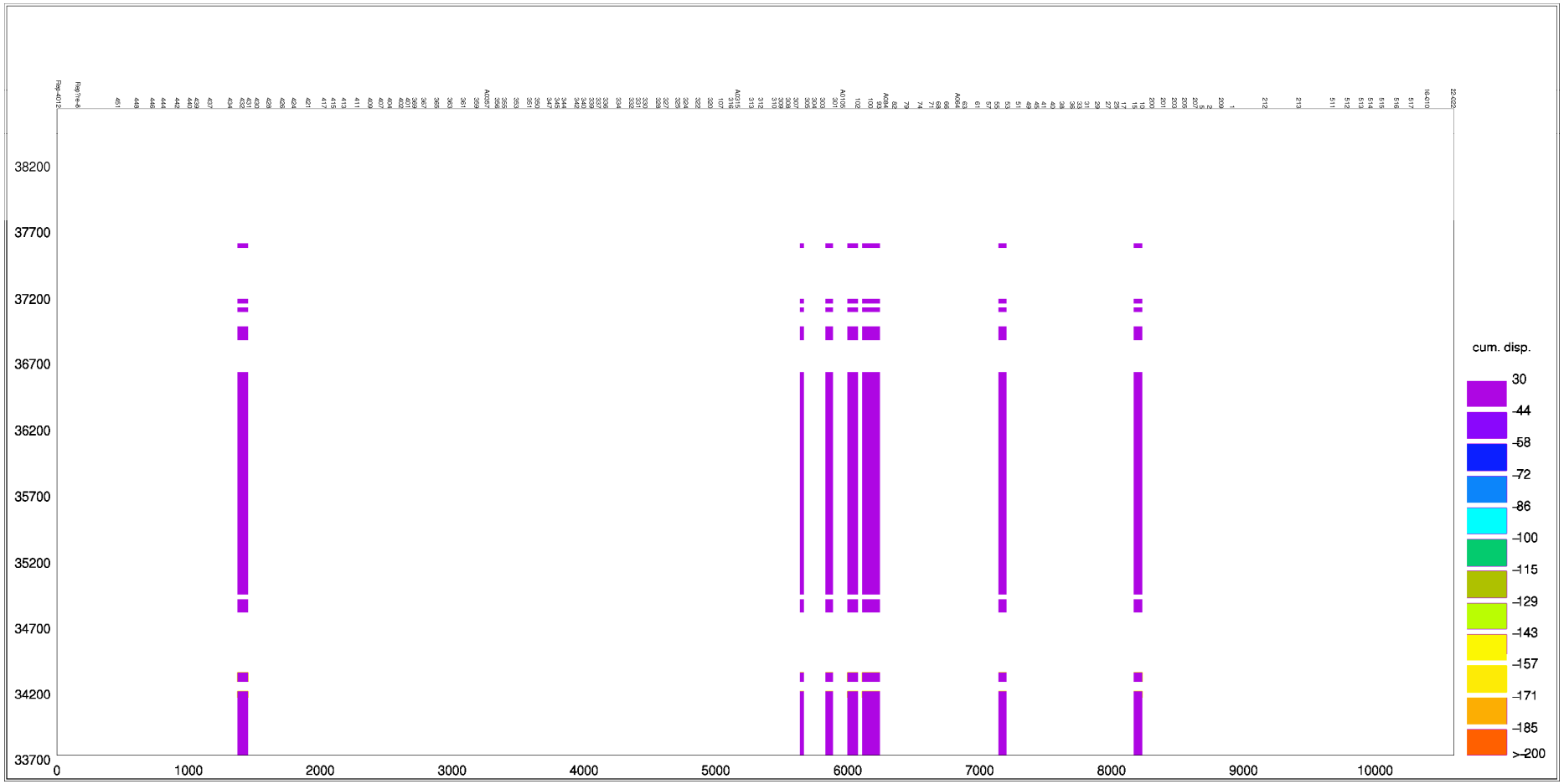


Figure 4.3.16 – Spatio-temporal profile of the PS-InSAR observations of processing chain 6, interpolated along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

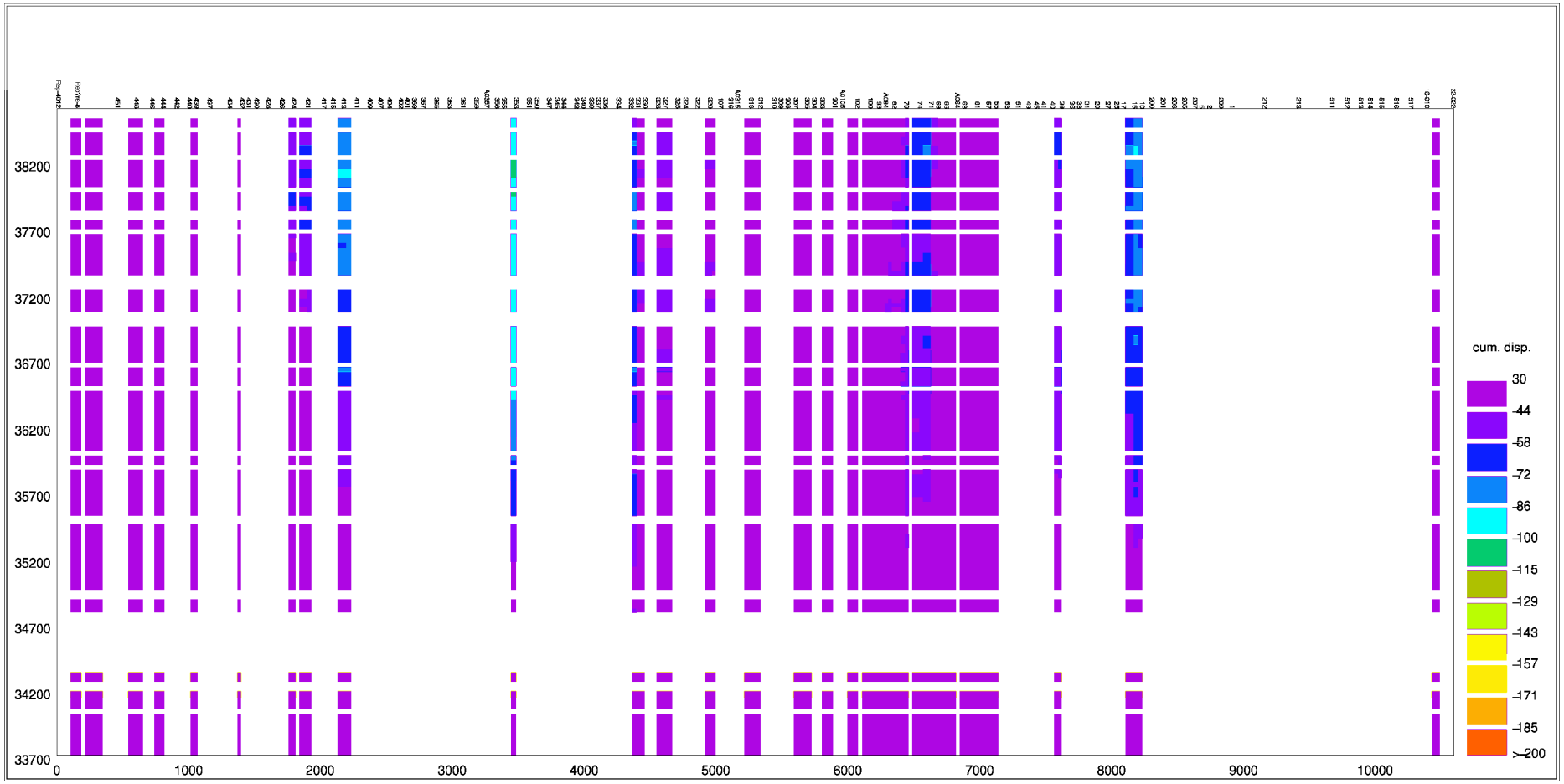


Figure 4.3.17 – Spatio-temporal profile of the PS-InSAR observations of processing chain 7, interpolated along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

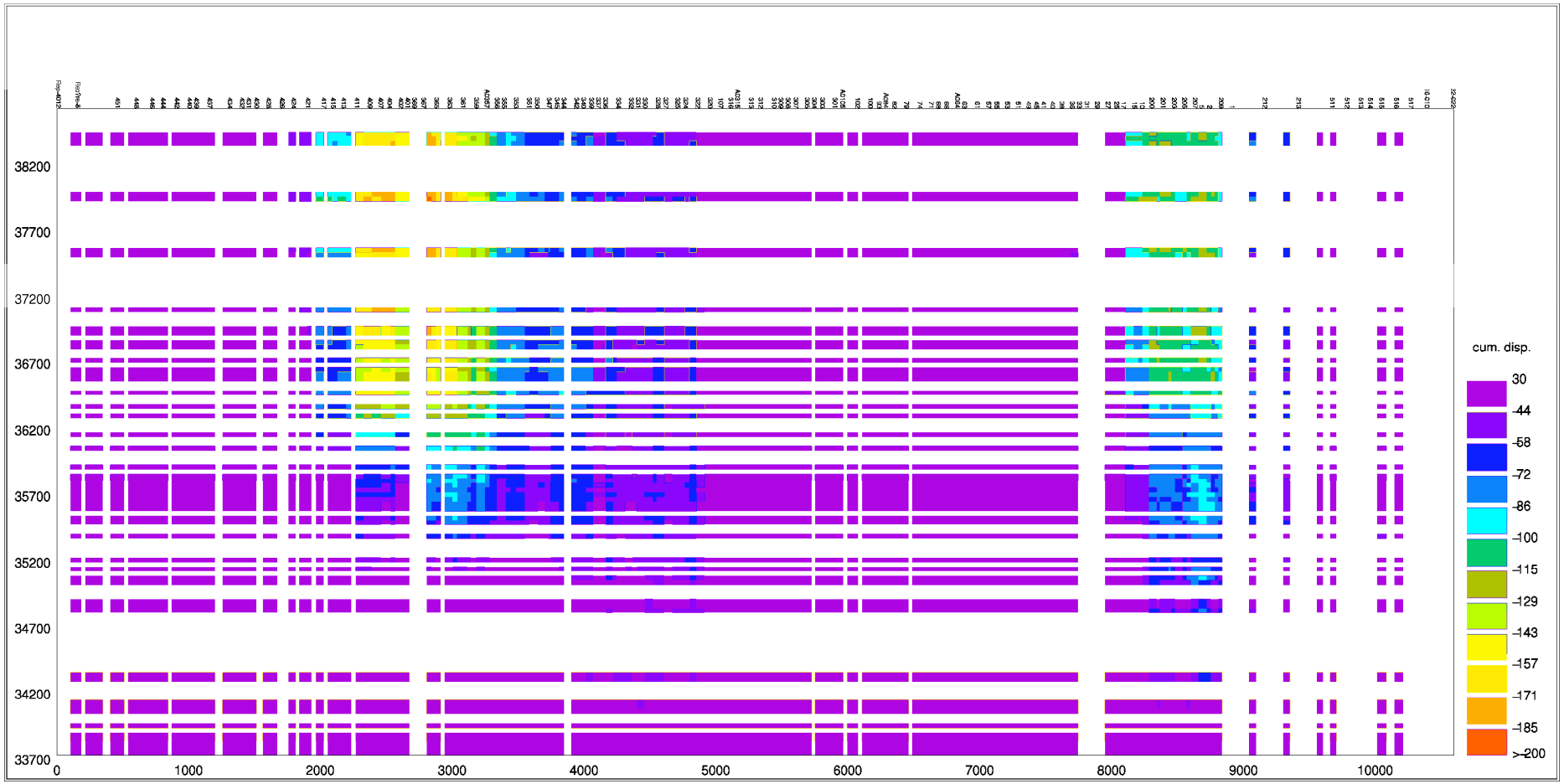


Figure 4.3.18 – Spatio-temporal profile of the PS-InSAR observations of processing chain 8, interpolated along levelling section SNCF (cumulative displacement in mm; horizontal lower axis: distance in m;horizontal upper axis: codes of levelling points; vertical axis: julian days (33800 = 15 july 1992)

Statistical Analysis

A statistical analysis has been performed in order to assess the goodness-of-fit of the PS-derived time series with respect to the time series obtained by the levelling campaigns. The following analyses have been made:

1. Regression analysis

2. Estimation of the Root Mean Squared Error of the PS-time series with respect to the levelling time series

3. Estimation of the goodness-of-fit of the PS-time series with respect to the levelling time series based on a non-parametric test. The *Kolmogorov-Smirnov*-test has been used.

4. Estimation of the goodness-of-fit of the PS-time series with respect to the levelling time series based on the *Student's-t*-test.

The data set consists of a total of 957 locations with a levelling time series. Of these 957 levelling time series there is a total of 368 time series starting before the first SAR-acquisition date, viz. 15 july 1992 and having a PS-InSAR time series from at least one processing chain. The other time series start after that date or have no associated PS-InSAR estimation. In order to have a common reference only those levelling time series have been used for further statistical analyses, having observations started before 15 july 1992.
The distribution of the levelling time series has been further characterized by the observed average velocity during the observation period. This is shown in table 4.2.1. Four of the five classes contain more than 70 levelling locations. A total of 17 levelling locations registrate an upward movement. The velocities are average velocities; at certain time intervals the velocity can be much larger. Table 4.2.1 also shows the number of time series with a maximum displacement larger than 1,4 cm (quarter of the SAR-wavelength) and 2,8 cm (half the SAR-wavelength) inbetween SAR-acquisition dates. As can be seen, up to an average velocity of -10m/yr, there are very little levelling time series showing a maximum displacement larger than half the radar-wavelength. Levelling time series with an average velocity smaller than -10 mm/yr however, in majority show maximum displacements larger than 2,8 cm. In total one third of the levelling time series show maximum displacements inbetween SAR-acquisition dates to be larger than 2,8 cm. About 75% of the levelling time series show maximum displacements of more than 1,4 cm inbetween SAR-acquisition dates.

velocity (mm/yr)	N	N-displacement > 1,4 cm	N-displacement > 2,8 cm
>0	17	16	1
0 - -5	96	20	0
-5 - -10	85	67	4
-10 - -15	71	61	45
<-15	99	99	74

Table 4.2.I – Number of levelling time series per velocity class and number of levelling time series with displacements larger than 2,8 cm inbetween SAR-acquisition dates

Table 4.2.2 shows the number of levelling time series which can be compared with PS-time series per processing chain. The total number varies beween 52 and 358 locations. As can be seen in table 2, however, 6 of the processing chains show an undersampling of levelling time series with maximum displacements larger than 2,8 cm inbetween acquisition dates. For each levelling time series the correlation coefficient, R^2, of the cumulative displacement observed by levelling and by PS-InSAR has been computed. A temporal resampling of the levelling time series has been used in order to have levelling observations concurrent at the SAR-acquisition dates. For most samples a positive correlation was obtained, meaning that the levelling time series and the PS-InSAR time series have a common trend. Weak correlation coefficients or negative correlation coefficients are seen for levelling time series with small negative or positive cumulative displacements. In one case, processing chain 6, predominantly negative correlation coefficients were obtained.

	N	N pos. correl.	N-displacement > 2,8 cm
1	52	30	16
2	180	146	51
3	120	81	23
4	147	87	29
5	142	110	38
6	53	3	2
7	145	133	36
8	358	305	120

Table 4.2.II – Number of estimated PS-time series colocated at levelling locations, number of positively correlated time series and number of time series with a maximum displacement larger than 2,8 cm inbetween SAR-acquisition dates.

Root Mean Squared Error (RMSE)

RMSE is a common statistic for quantifying elevation error (Shearer, 1990), and is used by the USGS as a data quality standard for many of its products, including DEMs (USGS, 1993). RMSE for a time series is defined by the equation

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(X_L(i) - X_{PS}(i)\right)^2} \qquad (9)$$

Here: N = number of paired observations
$X_L(i)$ = observation of levelling at date i
$X_{PS}(i)$ = observation of PS-InSAR at date i

RMSE represents the typical size of discrepancy between the two datasets at a certain levelling location, with values equaling or near zero indicating perfect or near perfect fit. The lower the RMSE, the better is the fit between the two-datasets.

Table 4.2.III shows the average RMSE per velocity class and per processing chain.

velocity (mm/yr)	1	2	3	4	5	6	7	8
>0	12	10	9	11	14	14	9	20
0 - -5	18	14	14	18	14	22	9	22
-5 - -10	38	41	38	46	32	58	25	27
-10 - -15	108	97	90	85	86	86	75	69
<-15	93	97	151	112	106	85	99	91

Table 4.2.III – Average RMSE per velocity class and per processing chain

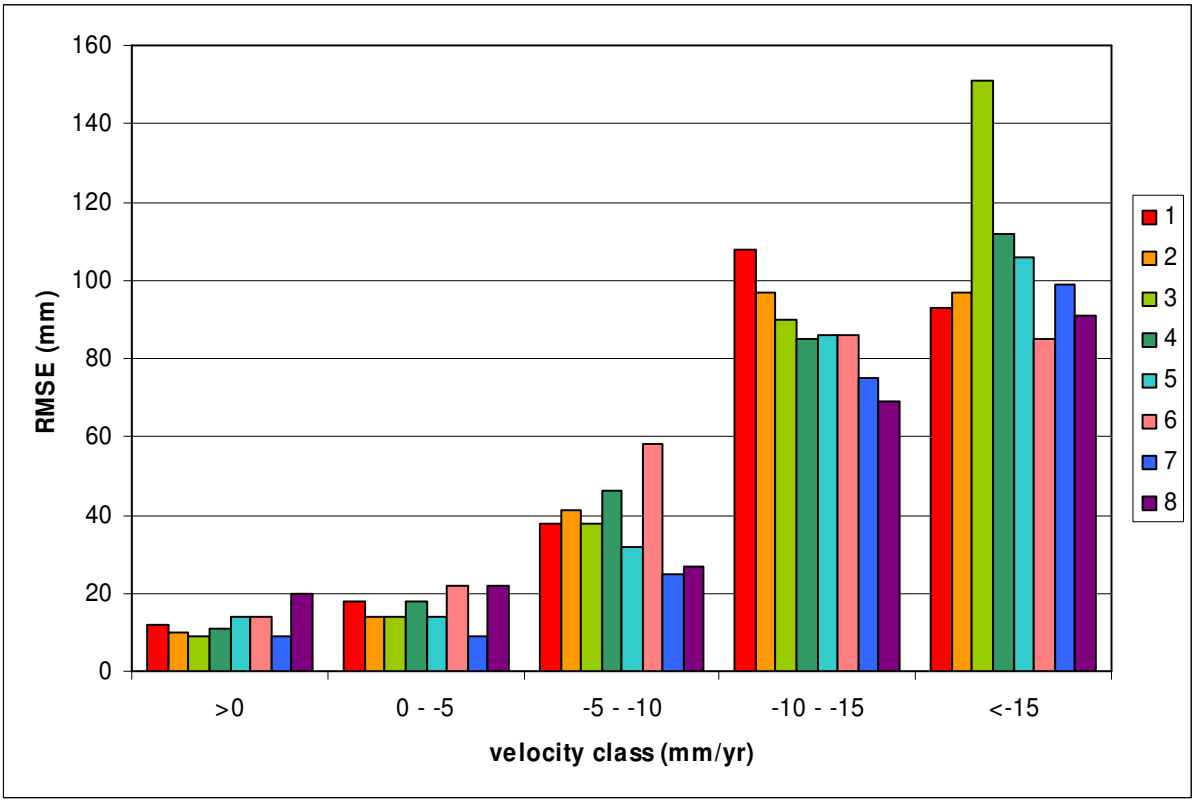


Figure 4.3.19 – Average RMSE per velocity class and per processing chain

As can be seen, lowest RMSE's are found for those time series which fall within the class -5 mm/yr - >0mm/yr. About two-third of the levelling time series of these velocity classes have a maximum displacement inbetween SAR-acquistion dates smaller than 1,4 cm. The average RMSE is about 10 – 15mm, apart from processing chain 8 which has an RMSE larger than 20 mm. RMSE increases considerably for those time series which fall in the class <-10 mm/yr.

Figure 4.3.20 shows for all processing chains the RMSE against the maximum displacement inbetween SAR-acquisition dates.

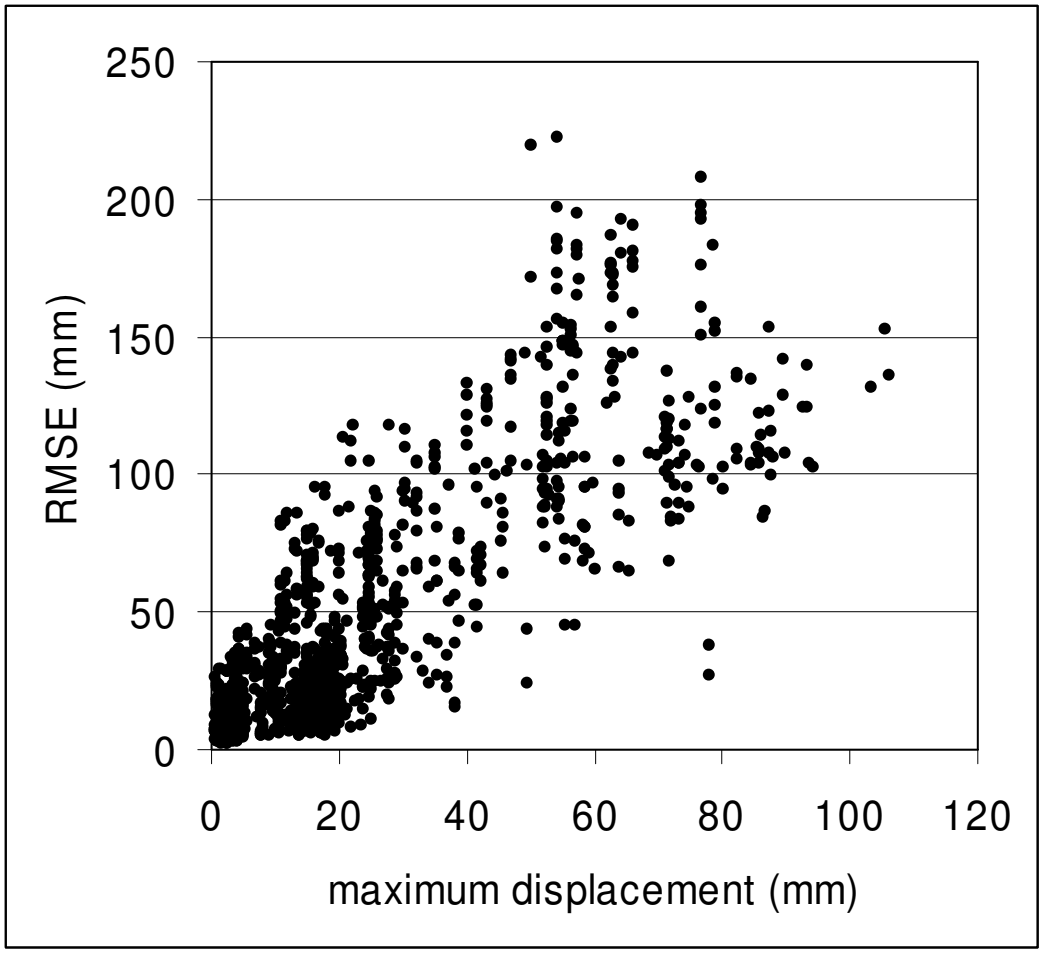


Figure 4.3.20 RMSE (in mm) against maximum displacementinbetween SAR-acquisition dates for all processing chains

The figure 4.3.20 shows that there is a correlation between RMSE and the maximum displacement inbetween SAR-acquisitions. This correlation is more pronounced for maximum displacements larger than about 20 mm.

Goodness of fit

All time series of all processing chains have been tested against the levelling time series using a KS-test and a students-t test. The test hypothesis is:

H_0: The mean of the diffeence of displacement of the PS-InSAR time series and the levelling time series is 0.
H_1: The mean of the diffeence of displacement of the PS-InSAR time series and the levelling time series is not 0.

These hypotheses are tested against a significance of 5% ($\alpha = 0{,}05$).

The KS-test resulted in ambiguous results, wrongly accepting or rejecting H_0 in a number of cases. This is due to the fact that for the time series comparison the

temporally resampled levelling time series have been used, resulting in a oversampling of certain displacement values.

As an alternative, the students-t test has been applied. The students-t test is a paired test which compares the test-statistic:

$$t = \frac{(\bar{x} - \mu) \cdot \sqrt{N}}{s}$$

Where: x = sample mean of the difference of displacement between PS-InSAR time series and levelling time series
μ = population mean of the difference of displacement between PS-InSAR time series and levelling time series (supposed to be 0)
N = number of observations per time series
s = sample standard deviation of the difference of the displacement between PS-InSAR time series and levelling time series

The t-value is compared whith $t_{\text{-test}}$ for a two-tailed students-t distribution for the number of observations. If $t < t_{\text{-test}}$ than H_0 is accepted at a significance of 0,05.

Figures *4.3.21, 4.3.22, 4.3.23* and table *4.3.IV* show the results for these computations. Four groups can be found. The first group consists of processing chains 1,3 and 6. These have very little to no time series tested positive for the H_0-hypothesis. Corrected for the number of time series having a displacement inbetween acquisition dates larger than *2,8* cm, these processing chains have 0 – 3% of their time series tested positive. Processing chains 5, 7 and 8 have more than 10% of the time series tested positive, if corrected for time series with a dsiplacement inbetween acquisition dates larger than *2,8* cm. Processing chains 2 and 4 are inbetween.

velocity (mm/yr)	1	2	3	4	5	6	7	8
>0	0	5	2	4	1	0	1	0
0 - -5	1	4	0	2	8	0	9	9
-5 - -10	0	2	0	2	2	0	2	10
-10 - -15	0	0	0	0	0	0	0	6
<-15	0	0	0	0	0	0	0	0

Table 4.3.IV – Number of time series for which the H_0-hypothesis is accepted at 5% significance level.

Some distinction can also be made with respect to the velocity classes. Apart from processing chains 1, 3 and 6 which have a very low number of time series tested positive, processing chains 2, 4, 5 and 7 seem to be more succesfull in measuring time series with low velocities. This reflects the outcomes of the RMSE-computations. Processing chain 8 on the other hand also succefully measures displacement time series at locations with average subsidence velocities smaller than -10mm/yr.

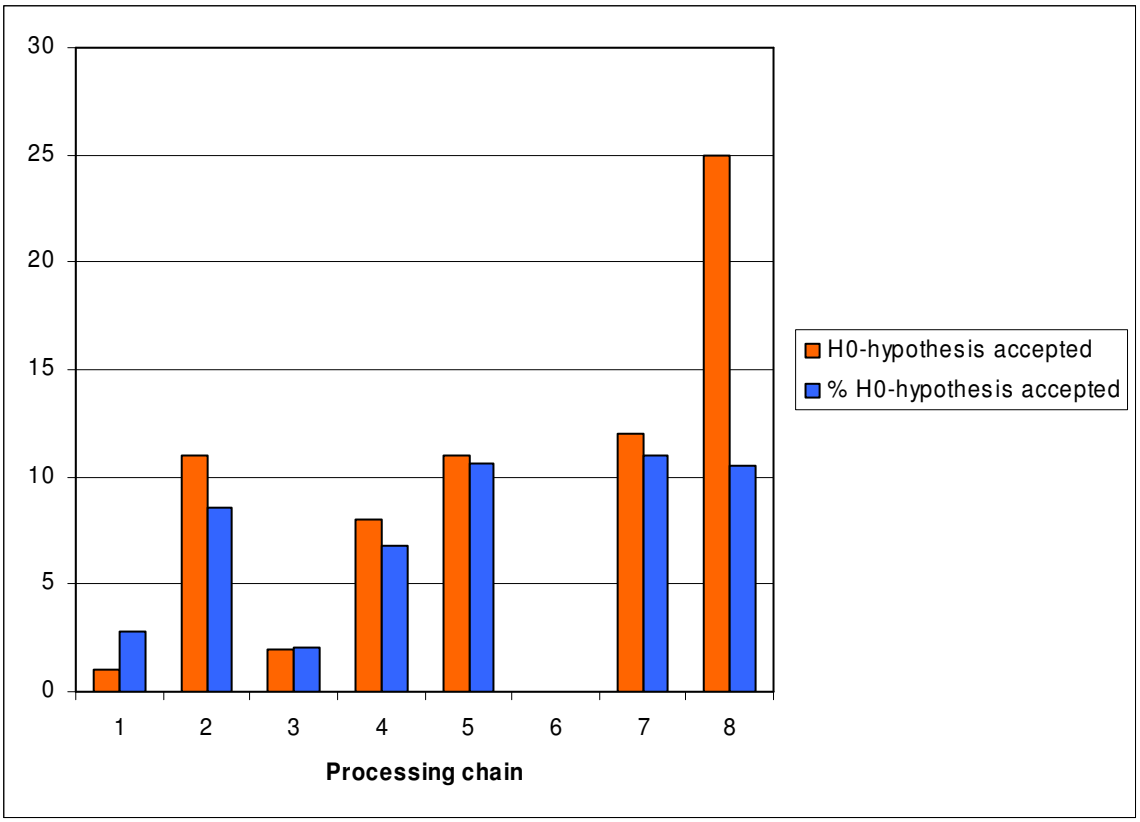


Figure 4.3.21 – Total number and percentage of time series for which the H_0-hypothesis has been accepted per processing chain

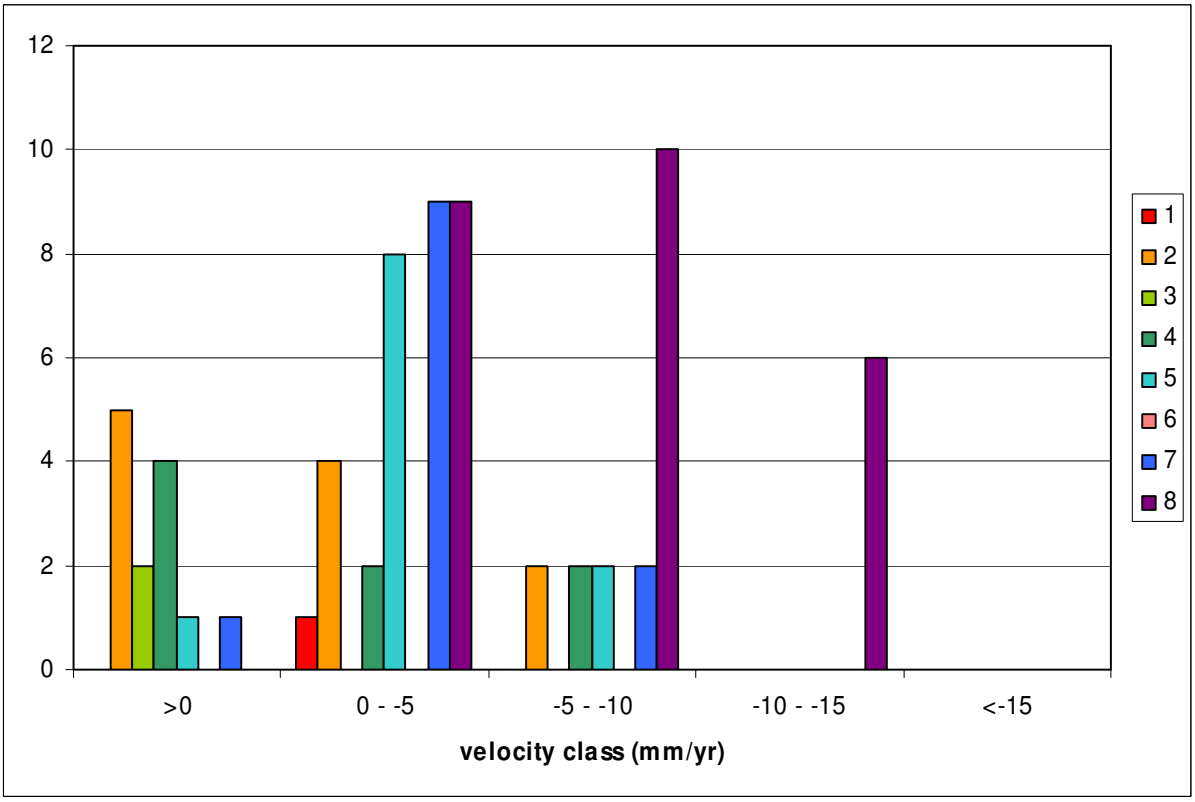


Figure 4.3.22 - Total number of time series for which the H_0-hypothesis has been accepted per velocity class (colour code refers to processing chain)

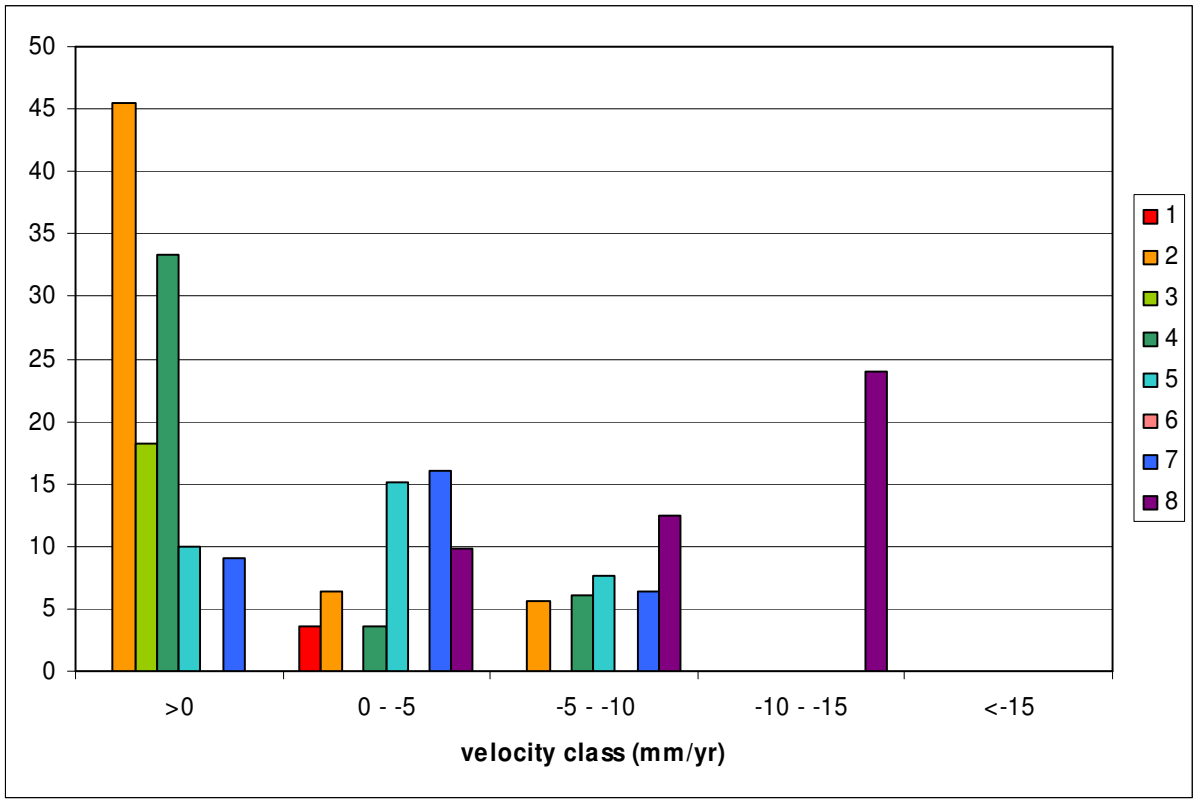


Figure 4.3.22 – Percentage of time series for which the H_0-hypothesis has been accepted per velocity class (colour code refers to processing chain)

4.4 Velocity validation

In the velocity validation task we used the temporally resampled levelling data (resampled at each SAR acquisition date) and the spatially resampled PSs data (resampled at each levelling point location) for comparison. In the figures below we show the maps of vertical velocities retrieved from cumulative displacement after resampling. We show vertical velocities detected by both the techniques at levelling locations and at SAR dates. We first show the vertical velocity map retrieved by levelling data by means of reference. The vertical velocity maps retrieved from PSs techniques for each team follow.

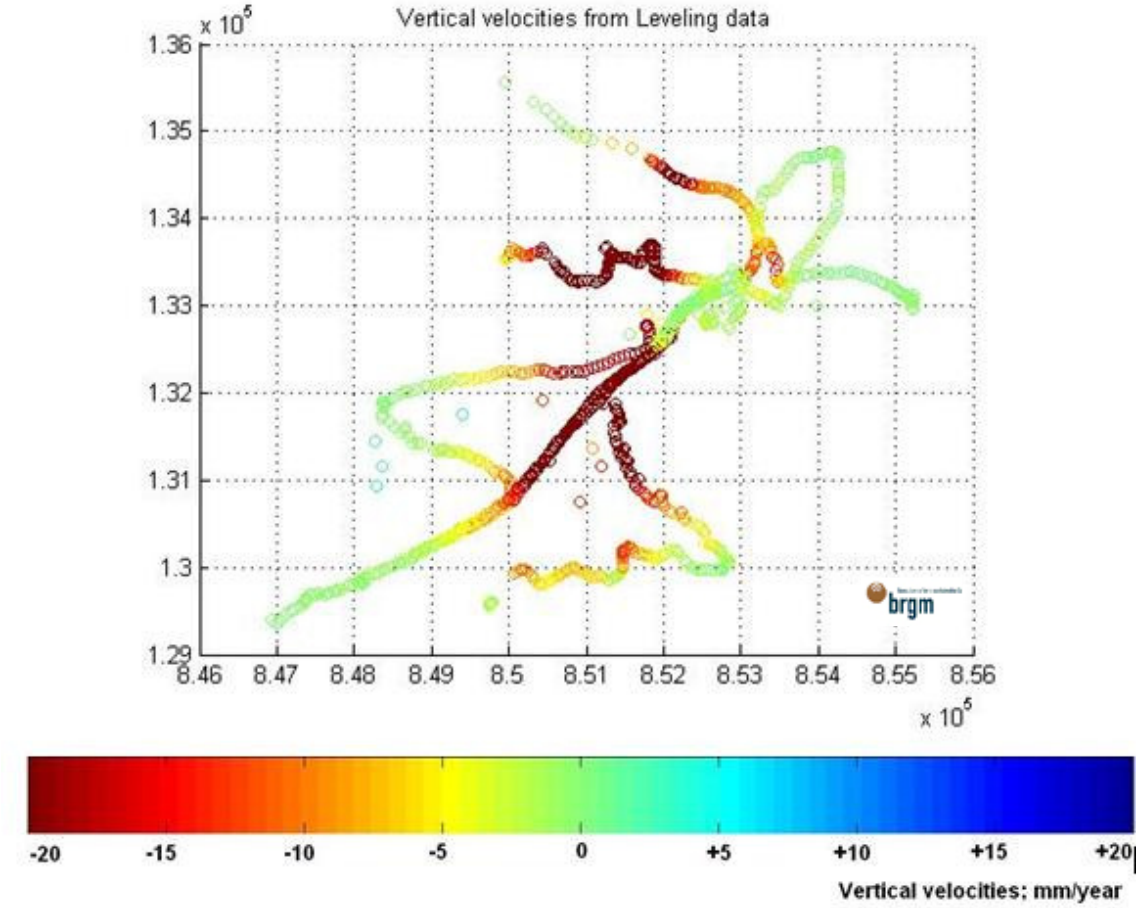


Figure 4.4.1: Vertical velocities from levelling data resampled at SAR acquisition dates

This figure has to be taken as comparison reference when looking at the figures in the next pages. It represents the map of vertical velocities retrieved by levelling measures time-resampled at SAR dates. In the next images, we show the vertical velocities retrieved by PSI techniques. In principle, both the technique should spot and measure the same rate of deformation. Although that, in this very study, we can see a certain number of differences.

Team 1

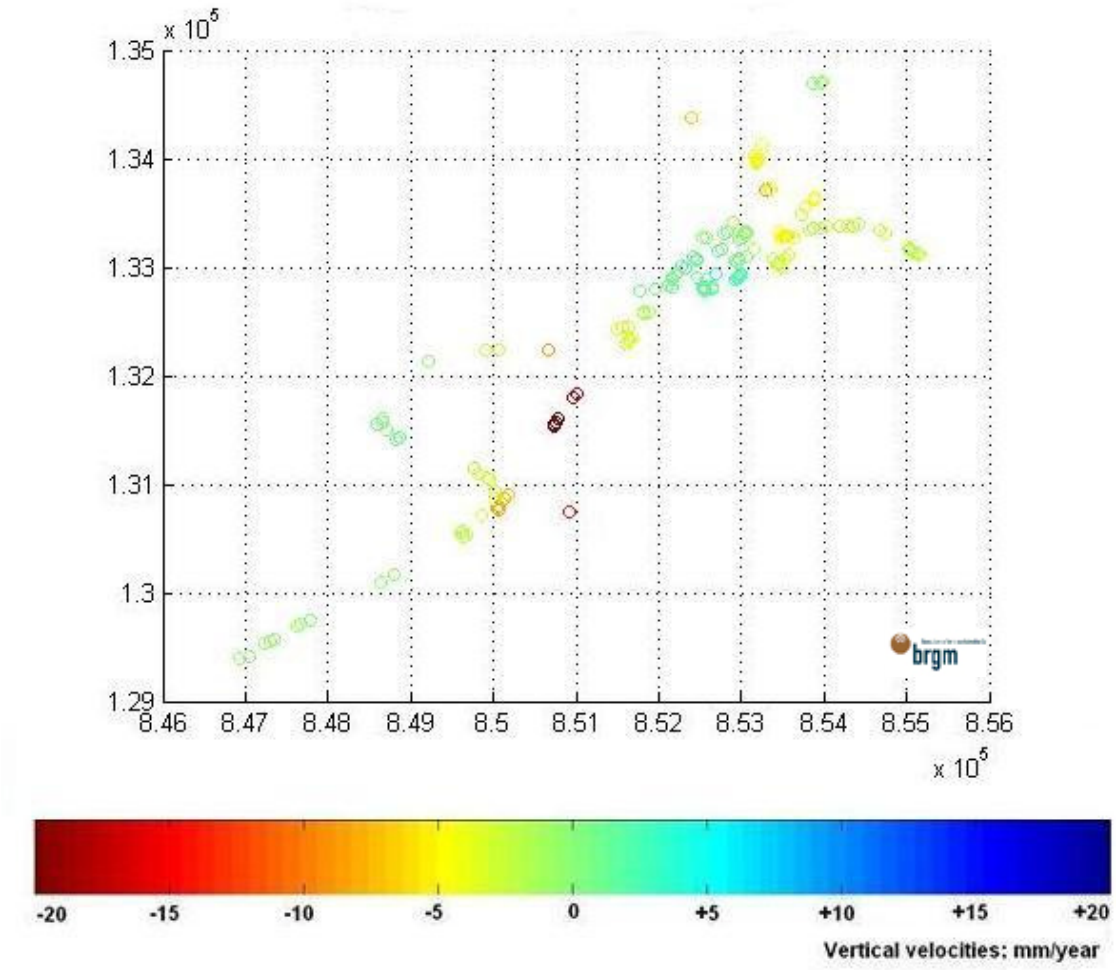


Figure 4.4.2: Vertical velocities of PSs resampled at levelling points location. Team 1.

A first visual assessment shows that the spatial resampling of Team 1 PS data produced a map of velocities which is quite similar to the map plotted from levelling data. The main subsidence areas are located, as well as the main stable areas. The individual report describes more accurately agreements and disagreements of Team 1 results with the levelling data for each specific area.

It should be noticed that Team 1 suffers from the lack of data being comparable with levelling data. However, the outcome shows good agreement with the levelling data in stable areas, although they are sometimes identified as small uplifts. This leads to errors of up to 4mm/year. When data exist, the subsidence areas are located, but are underestimated (up to 5mm/year).

Team 2

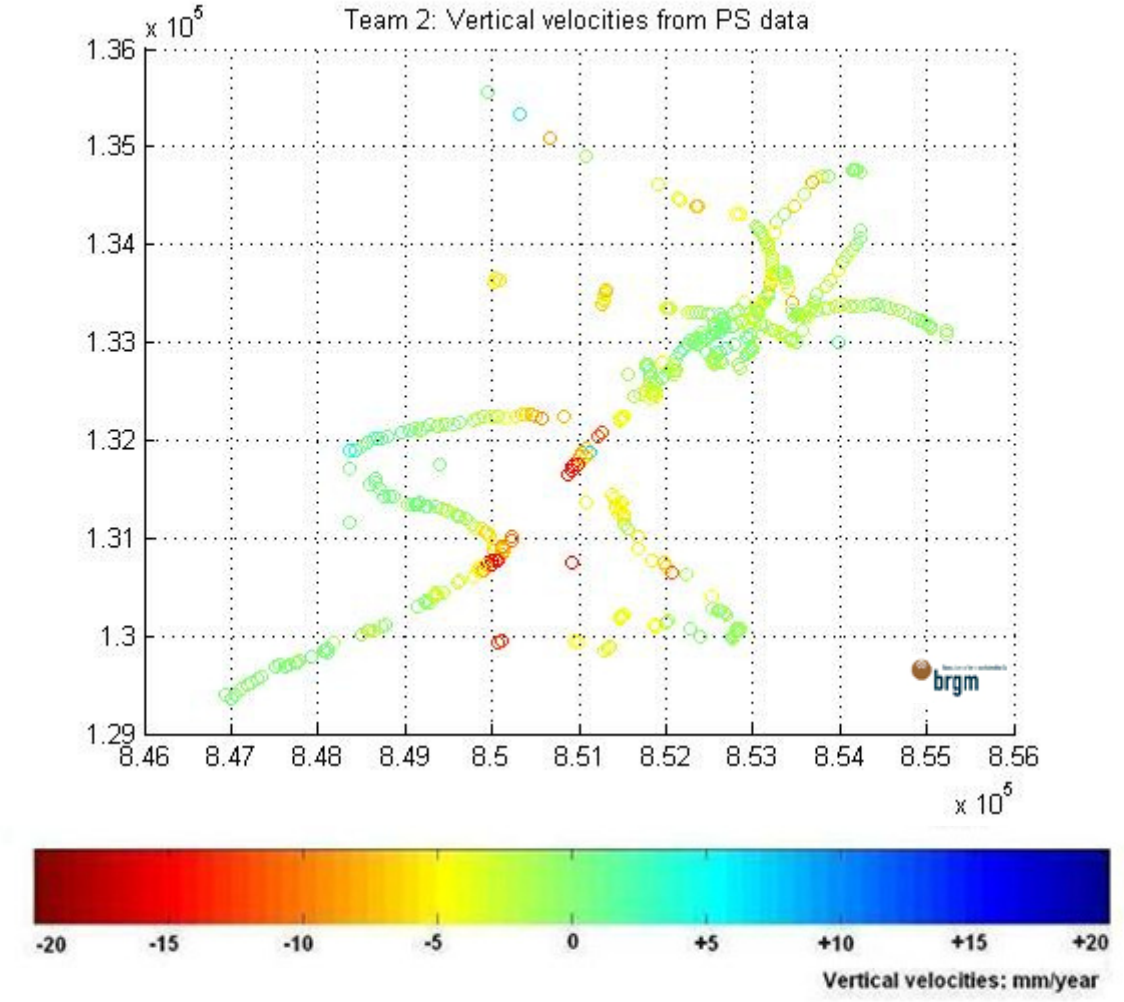


Figure 4.4.3: Vertical velocities of PSs resampled at levelling points location. Team 2.

A first visual assessment shows that the spatial resampling of Team 2 PS data produced a map of velocities which outlines the main trends that appear on the map plotted from the levelling data. The main subsidence areas are located, as well as the main stable areas. The individual report describes more accurately agreements and disagreements of Team 2 results with the levelling data for each specific area.

Team 2 data shows good agreement with the levelling data for the stable areas. However, these areas are sometimes identified as small uplifts. In these cases, disagreements are of the order of 3-4 mm/year at the most. The subsidence areas are located, but are underestimated. Finally, no PS data could be resampled for subsidence areas with a maximum velocity.

Team 3

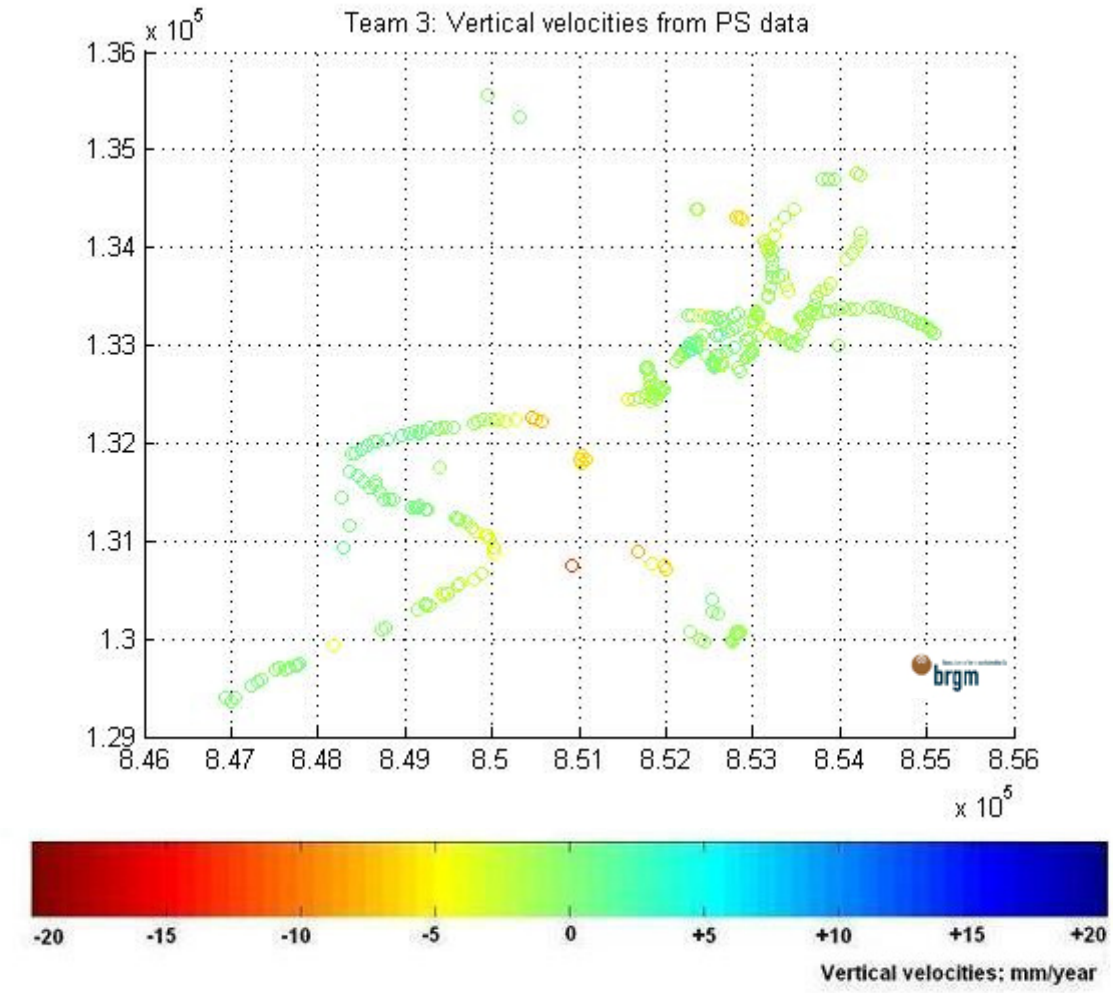


Figure 4.4.4: Vertical velocities of PSs resampled at levelling points location. Team 3.

A first visual assessment shows that the spatial resampling of Team 3 PS data produced a map of velocities which outlines somehow the main trends that appear on the map plotted from the levelling data. The main subsidence areas are located, as well as the main stable areas. The individual report describes more accurately agreements and disagreements of Team 3 results with the levelling data for each specific area.

Team 3 data shows good agreement with the levelling data for the stable areas. However, these areas are sometimes identified as small uplifts. In these cases, disagreements are of the order of 2mm/year. The subsidence areas are most of the time not located, and if located disagree with the levelling data (scale of error: more than 10mm/year). Finally, no PS data could be resampled for subsidence areas with a maximum velocity.

Team 4

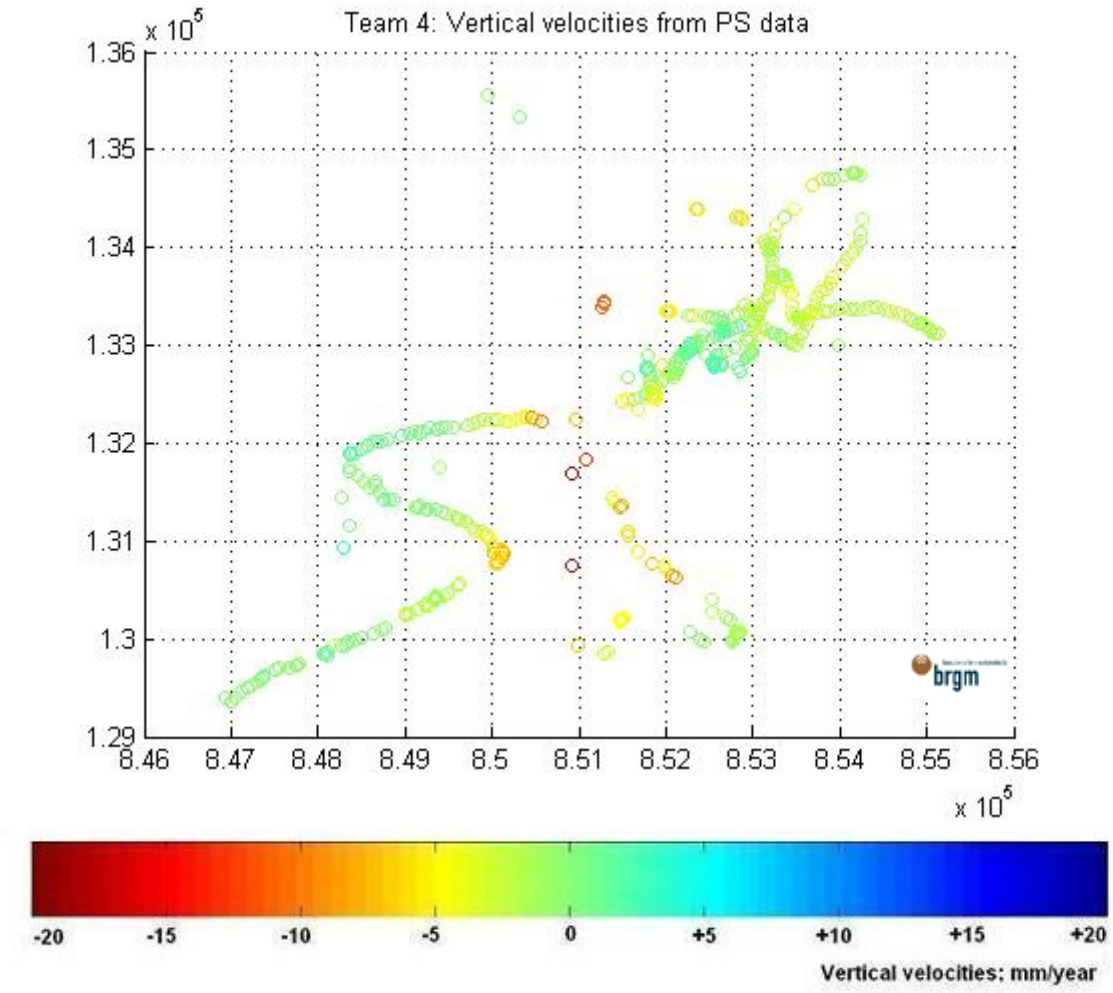


Figure 4.4.5 Vertical velocities of PSs resampled at levelling points location. Team 4.

A first visual assessment shows that the spatial resampling of Team 4 PS data produced a map of velocities which outlines the main trends that appear on the map plotted from the levelling data. The main subsidence areas are located, as well as the main stable areas. The individual report describes more accurately agreements and disagreements of Team 4 results with the levelling data for each specific area.

Team 4 data shows good agreement with the levelling data for the stable areas. However, these areas are sometimes identified as small uplifts. In these cases, disagreement is of the order of 2mm/year. The subsidence areas are hardly identified, and if identified, strongly disagree with the levelling data (scale of error: more than 10mm/year). Finally, no PS data could be resampled for subsidence areas with a maximum velocity.

Team 5

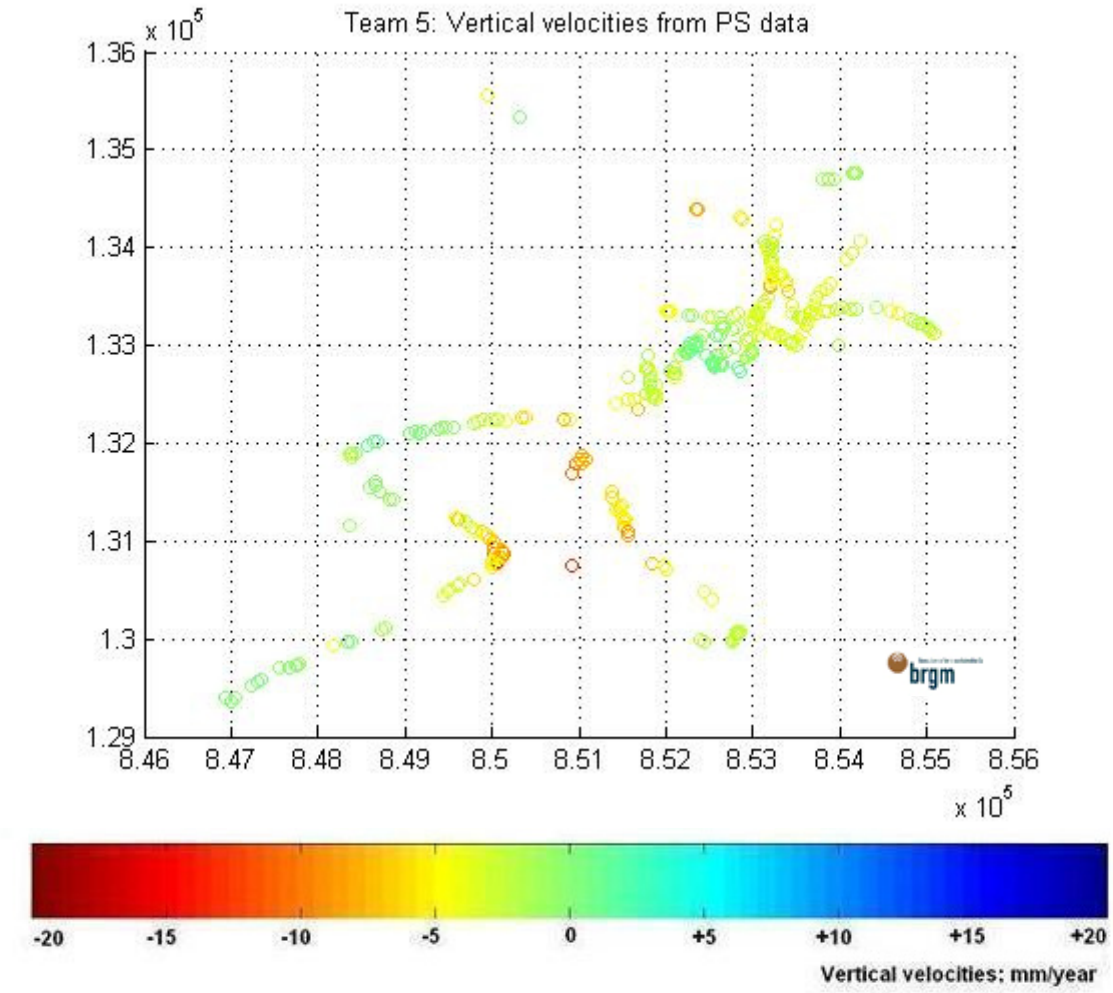


Figure 4.4.6: Vertical velocities of PSs resampled at levelling points location. Team 5.

A first visual assessment shows that the spatial resampling of Team 5 PS data produced a map of velocities which outlines the main trends that appear on the map plotted from the levelling data. The main subsidence areas are located, as well as the main stable areas. The individual report describes more accurately agreements and disagreements of Team 5 results with the levelling data for each specific area.

Team 5 data shows good agreement with the levelling data for the stable areas, but with a low accuracy of about 2-3mm/year. When data exist, the subsidence areas are identified, but are underestimated and strongly disagree with the levelling data (scale of error: more than 10mm/year). Finally, no PS data could be resampled for subsidence areas with a maximum velocity.

Team 6

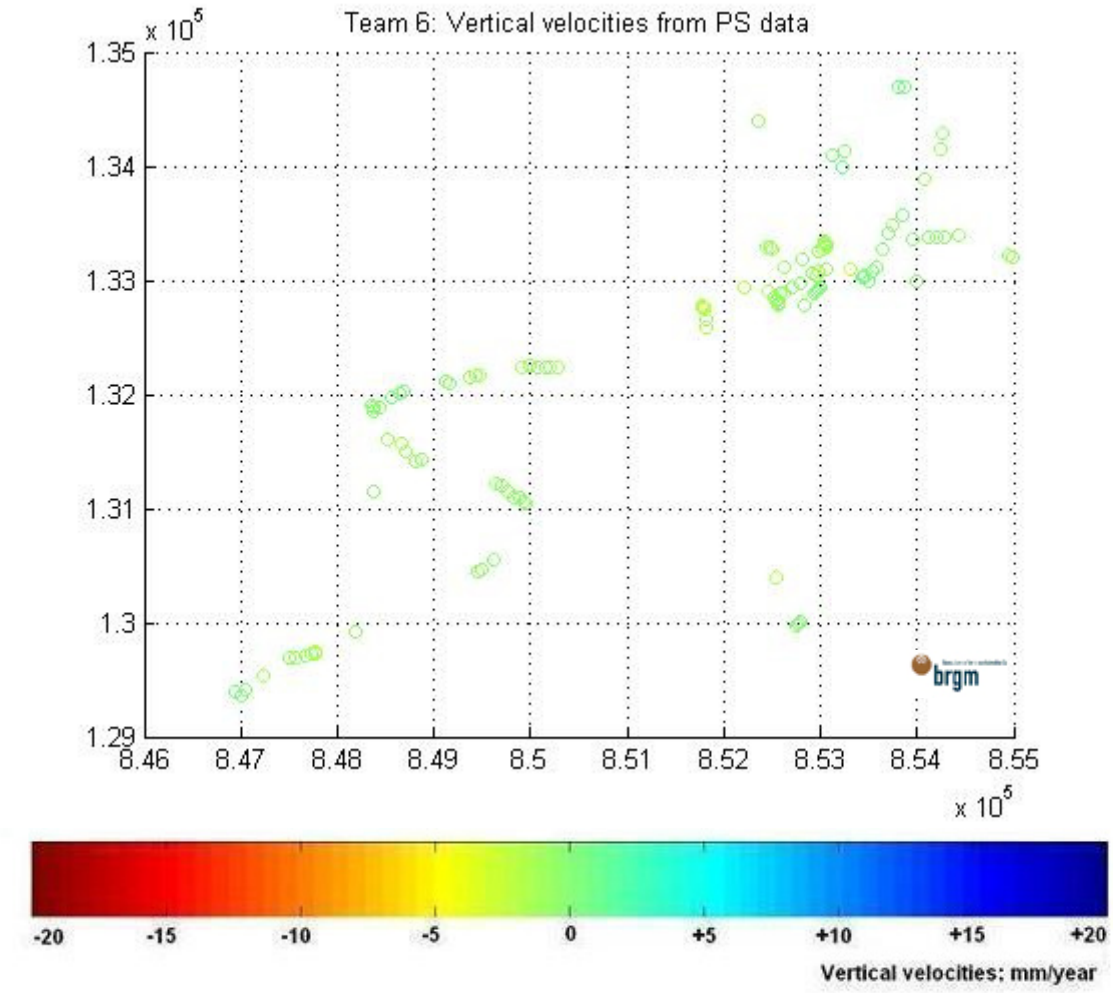


Figure 4.4.7: Vertical velocities of PSs resampled at levelling points location. Team 6.

A first visual assessment shows that the spatial resampling of Team 6 PS data produced a map of velocities showing a uniform stable area.

Team 6 PS data allow identifying the stable areas, though no PS data were found in the subsidence areas. The velocities disagree with the levelling data of about 2-3mm/year in stable areas.

Team 7

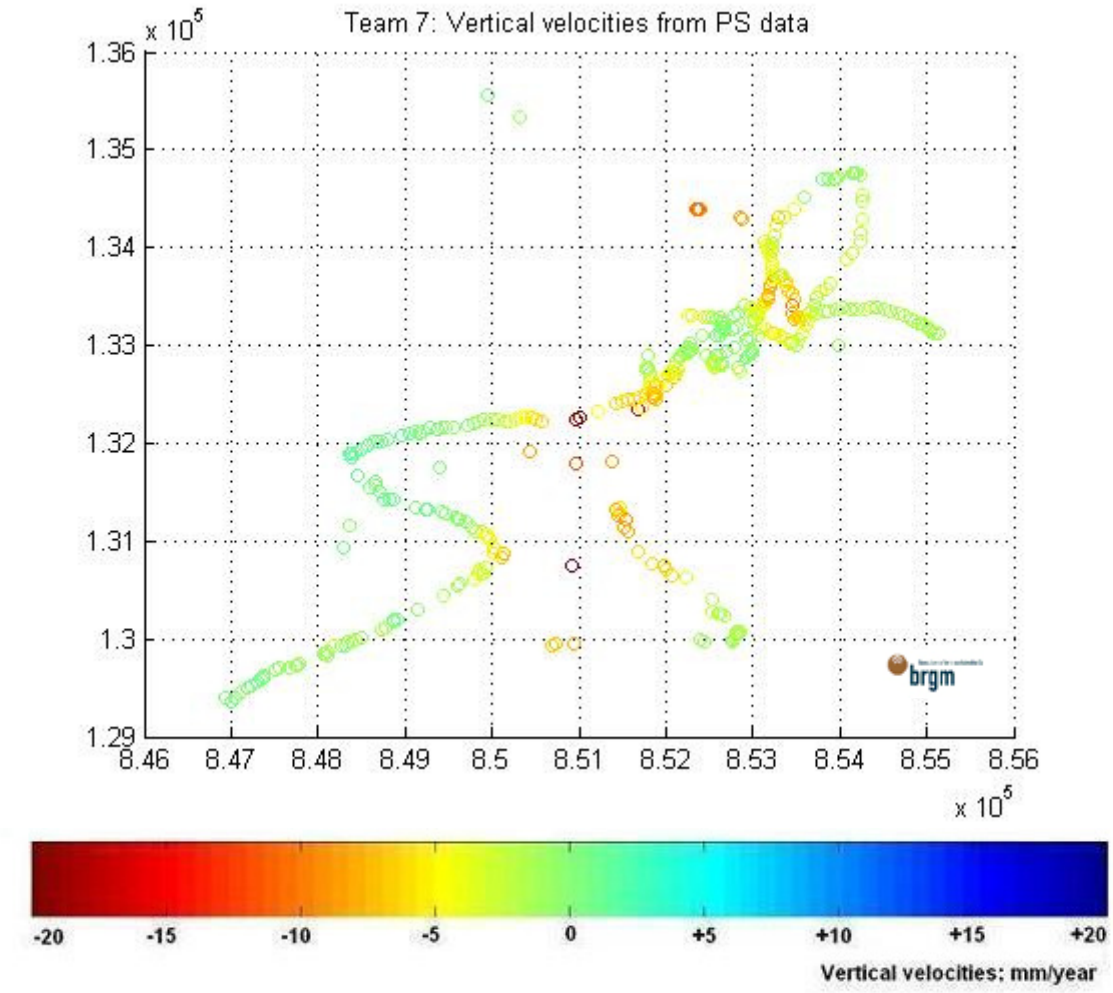


Figure 4.4.8: Vertical velocities of PSs resampled at levelling points location. Team 7.

A first visual assessment shows that the spatial resampling of Team 7 PS data produced a map of velocities which outlines in some way the main trends that appear on the map plotted from the levelling data. The main subsidence areas are more or less located, as well as the main stable areas. The individual report describes more accurately agreements and disagreements of Team 7 results with the levelling data for each specific area.

Team 7 suffers from the lack of data being comparable with levelling data after the resampling procedure. These data allow identifying the stable areas, but too few PS data were found in the subsidence areas. The retrieved velocities disagree with the levelling data of about 2-3mm/year in stable areas. No PS data could be resampled in subsidence areas.

Team 8

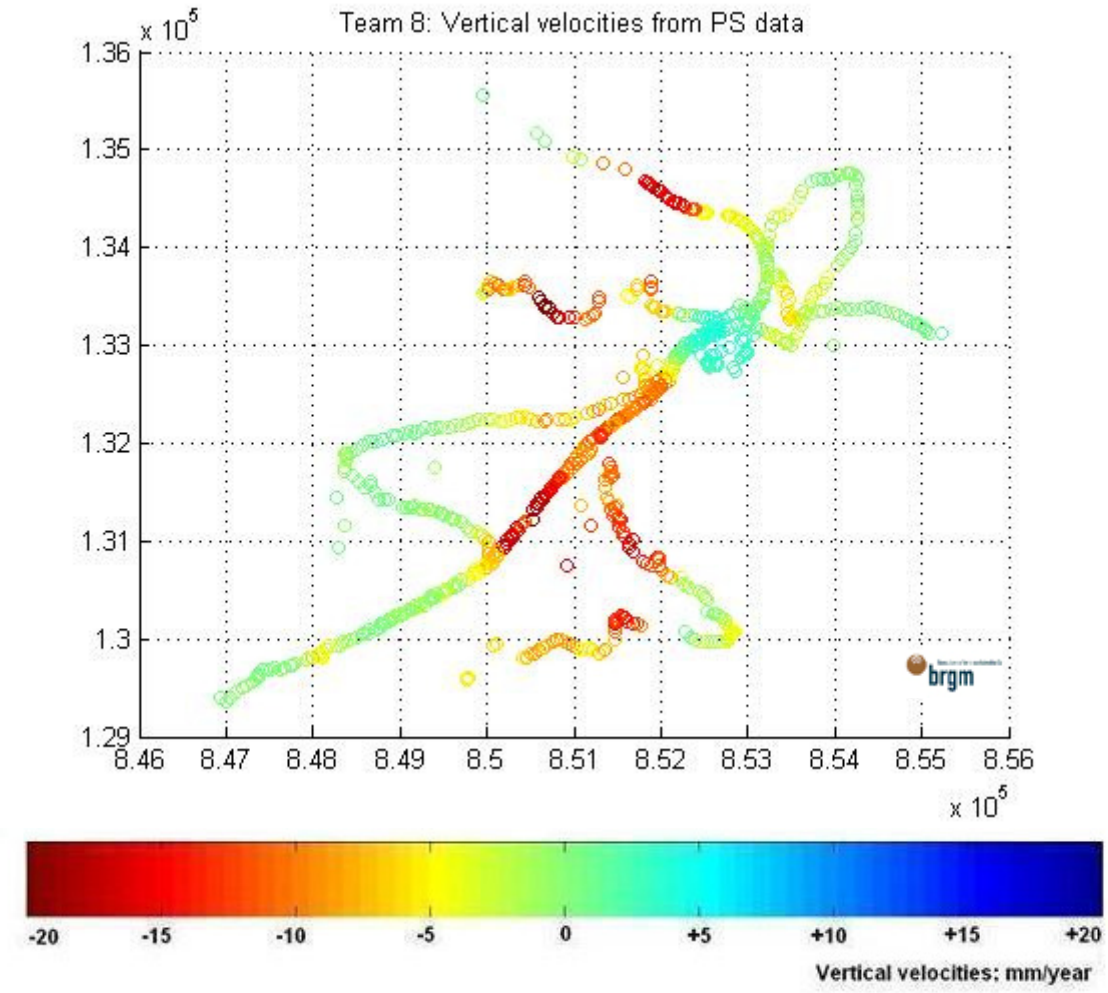


Figure 4.4.9: Vertical velocities of PSs resampled at levelling points location. Team 8.

A first visual assessment shows that the spatial resampling of Team 8 PS data produced a map of velocities which is very similar to the map retrieved from levelling data. The subsidence and stable areas are located. The individual report describes more accurately agreements and disagreements of Team 8 results with the levelling data for each specific area.

Team 8 data shows good agreement with the levelling data for the stable areas, sometimes with a lower accuracy of about 2-3mm/year. The subsidence areas are well identified, but are underestimated (scale of error: 10mm/year). It should be noted that PS data could be resampled for subsidence areas with a maximum velocity.

In the next paragraph we are showing and quantifying the difference between deformations measured by levelling and by PSI techniques. Rates of deformation have been put into classes and compared in terms of values and number of points within each class.

Summary of visual comparison of estimated velocities

In this paragraph we summarize differences and similarities between vertical velocities retrieved from levelling values and from PS values. The figure below (4. 4.10) illustrates the locations of the main subsidence area as well as the main stable area in Gardanne.

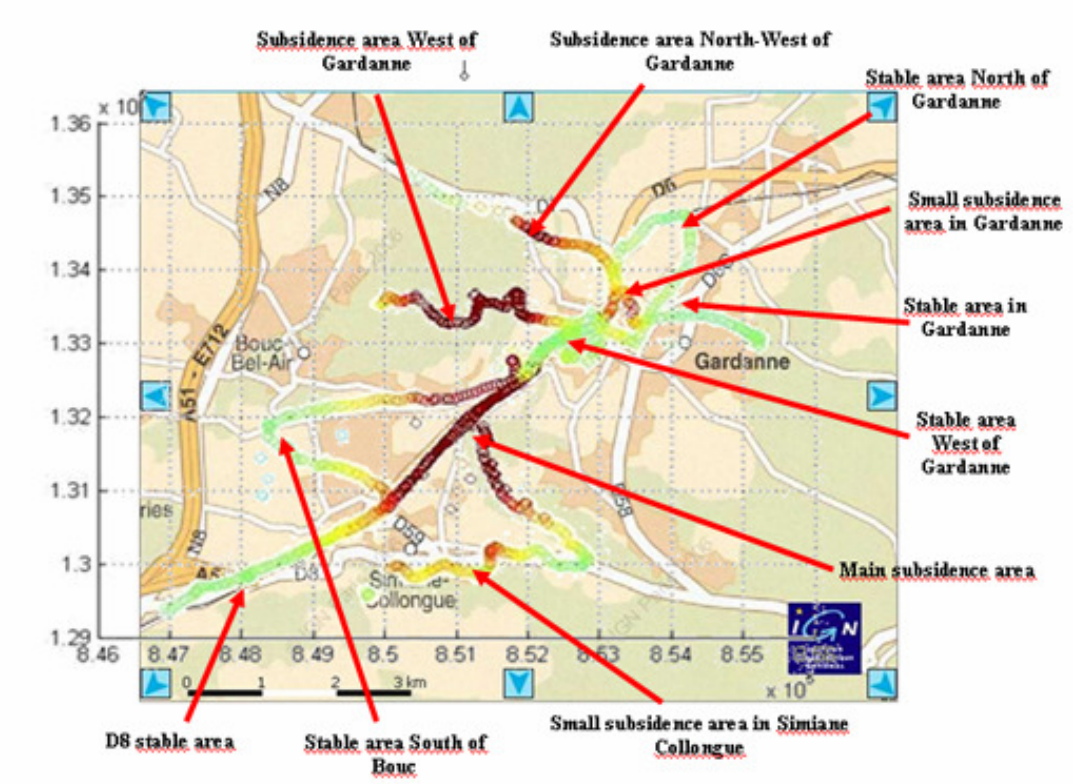


Figure 4.4.10. Map of Gardanne showing position of subsidence areas as identified from the levelling data

Area	Displacements as identified by PS	Scale of disagreement
Main subsidence area	Identified, but underestimated	5mm/year
Subsidence area north west from Gardanne	No data	
Subsidence area west from Gardanne	No data	
Small subsidence area in Gardanne	Too few data	
Small subsidence area in Simiane Collongue	No data	
Stable area north from Gardanne	Identified as small uplift (only a few data available)	3-4 mm/year
Stable area west from Gardanne	No data	
Stable area in Gardanne	Identified	Good agreement
Stable area south from Bouc	Identified but too few data	Good agreement
D8 stable area	Identified as small uplift (only a few data available)	2-3 mm/year

Table 4.4.I - TEAM 1

Area	Displacements as identified by PS	Scale of disagreement
Main subsidence area	Identified, but underestimated	More than 10mm/year
Subsidence area north west from Gardanne	Identified, but underestimated	5mm/year
Subsidence area west from Gardanne	Identified, but underestimated	More than 10mm/year
Small subsidence area in Gardanne	Identified, but underestimated	More than 10mm/year
Small subsidence area in Simiane Collongue	Identified (only a few data available)	Good agreement for most of the points
Stable area north from Gardanne	Identified as small uplift	3-4 mm/year
Stable area west from Gardanne	No data	
Stable area in Gardanne	Identified as a small uplift	2 mm/year
Stable area south from Bouc	Identified as a small uplift	2mm/year
D8 stable area	Identified	Good agreement

Table 4.4.II - TEAM 2

Area	Displacements as identified by PS	Scale of disagreement
Main subsidence area	Identified, but underestimated (only a few data available)	More than 10mm/year
Subsidence area north west from Gardanne	Too few data	
Subsidence area west from Gardanne	No data	
Small subsidence area in Gardanne	Not identified	
Small subsidence area in Simiane Collongue	No data	
Stable area north from Gardanne	Identified	Good agreement
Stable area west from Gardanne	Identified as small uplift	About 2 mm/year
Stable area in Gardanne	Identified	Good agreement
Stable area south from Bouc	Identified as a small uplift	2mm/year
D8 stable area	Identified	Good agreement

Table 4.4.III - TEAM 3

Area	Displacements as identified by PS	Scale of disagreement
Main subsidence area	Too few data	
Subsidence area north west from Gardanne	Identified, but underestimated (only a few data available; could not find PS in maximum displacements areas)	More than 10mm/year
Subsidence area west from Gardanne	Identified, but underestimated (only a few data available)	More than 10mm/year
Small subsidence area in Gardanne	Identified, but underestimated	More than 10mm/year
Small subsidence area in Simiane Collongue	Too few data	
Stable area north from Gardanne	Identified	Good agreement
Stable area west from Gardanne	Identified as small uplift	About 2 mm/year
Stable area in Gardanne	Identified	Good agreement
Stable area south from Bouc	Identified as a small uplift	2mm/year
D8 stable area	Identified as a small uplift	2mm/year

Table 4.4.IV - TEAM 4

Area	Displacements as identified by PS	Scale of disagreement
Main subsidence area	Identified, but underestimated (only a few data available; could not find PS in maximum displacements areas)	More than 10mm/year
Subsidence area north west from Gardanne	Too few data	
Subsidence area west from Gardanne	Too few data	
Small subsidence area in Gardanne	Identified, but underestimated	More than 10mm/year
Small subsidence area in Simiane Collongue	Too few data	
Stable area north from Gardanne	Identified (only a few data available)	2-3mm/year
Stable area west from Gardanne	Identified	About 2 mm/year
Stable area in Gardanne	Identified as a small subsidence	About 2 mm/year
Stable area south from Bouc	Identified	2-3mm/year
D8 stable area	Identified	Good agreement

Table 4.4.V - TEAM 5

Area	Displacements as identified by PS	Scale of disagreement
Main subsidence area	Too few data	More than 10mm/year for the few existing PS
Subsidence area north west from Gardanne	Too few data	
Subsidence area west from Gardanne	Too few data	
Small subsidence area in Gardanne	Too few data	
Small subsidence area in Simiane Collongue	Too few data	
Stable area north from Gardanne	Identified as small uplift (only a few data available)	2-3mm/year
Stable area west from Gardanne	Identified as small uplift	2-3mm/year
Stable area in Gardanne	Identified as small uplift	2-3mm/year
Stable area south from Bouc	Identified as small uplift	2-3mm/year
D8 stable area	Identified as small uplift	2-3mm/year

Table 4.4.VI - TEAM 6

Area	Displacements as identified by PS	Scale of disagreement
Main subsidence area	Too few data	
Subsidence area north west from Gardanne	Too few data	
Subsidence area west from Gardanne	Too few data	
Small subsidence area in Gardanne	Too few data	
Small subsidence area in Simiane Collongue	Too few data	
Stable area north from Gardanne	Identified (only a few data available)	2-3mm/year
Stable area west from Gardanne	Identified	2-3mm/year
Stable area in Gardanne	Identified	2-3mm/year
Stable area south from Bouc	Identified as small uplift	2-3mm/year
D8 stable area	Identified	2 mm/year

Table 4.4.VII - TEAM 7

Area	Displacements as identified by PS	Scale of disagreement
Main subsidence area	Identified, but underestimated	10 mm/year
Subsidence area north west from Gardanne	Identified, but underestimated	5 mm/year
Subsidence area west from Gardanne	Identified, but underestimated	10 mm/year
Small subsidence area in Gardanne	Identified, but underestimated. Partly identified as a stable area	10 mm/year
Small subsidence area in Simiane Collongue	Identified	2 mm/year
Stable area north from Gardanne	Identified	Good agreement
Stable area west from Gardanne	Identified as a small uplift	2-3mm/year
Stable area in Gardanne	Identified	Good agreement
Stable area south from Bouc	Identified	Good agreement
D8 stable area	Identified	Good agreement

Table 4.4.VIII - TEAM 8

4.4.1. Characterization of PS velocity consistency with levelling velocity

In order to characterize the disagreements of velocities between PS and levelling resampled data, we classified all levelling points upon the criteria shown in the table bellow. Then we classified PSs according to the same criteria and we compared the two. How does PSI technique perform over low, medium and high deformation rates? This comparison allows the assessing of PSI performance upon dependence of different deformation rates.

Class A +	Points whose velocities estimated from the temporally resampled levelling data is above 0mm/year
Class A -	Points whose velocities estimated from the temporally resampled levelling data is between -5 mm/year and 0 mm/year
Class B -	Points whose velocities estimated from the temporally resampled levelling data is between -10 mm/year and -5 mm/year
Class C -	Points whose velocities estimated from the temporally resampled levelling data is between -15 mm/year and -10 mm/year
Class D -	Points whose velocities estimated from the temporally resampled levelling data is under -15 mm/year

For each class, the mean and standard deviation of the estimated PS velocity for all points belonging to one given class was computed. It is compared with the mean and standard deviation of the estimated levelling velocity for each point belonging to the same class.

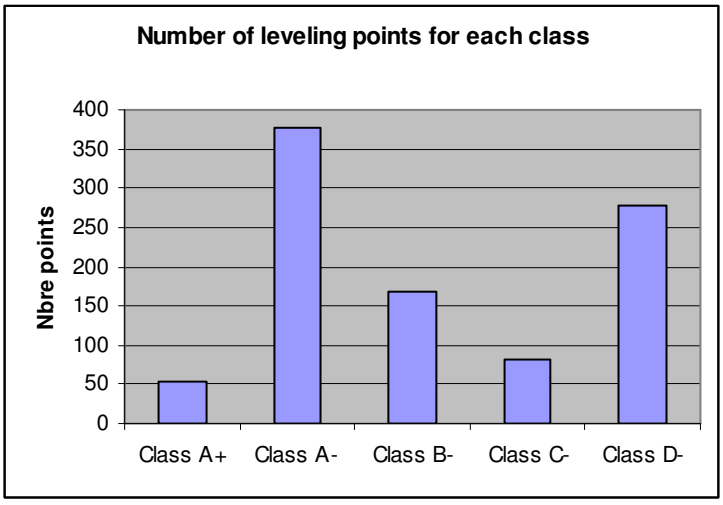


Figure 4.4.11: Number of levelling points for each class.

Team 1

Class	Mean velocities levelling (mm/year)	Standard deviation velocities levelling (mm/year)	Mean velocities PS (mm/year)	Standard deviation velocities PS (mm/year)	Number Points
All	-5.3	7.1	-2.5	4.5	158
A +	0.2	0.18	-1.7	1.5	13
A -	-1.3	1.2	-0.6	2.1	88
B -	-7.0	1.4	-3.1	2.0	29
C -	-12.0	1.5	-6.6	5.6	9
D -	-21.8	4.0	-8.6	8.4	19

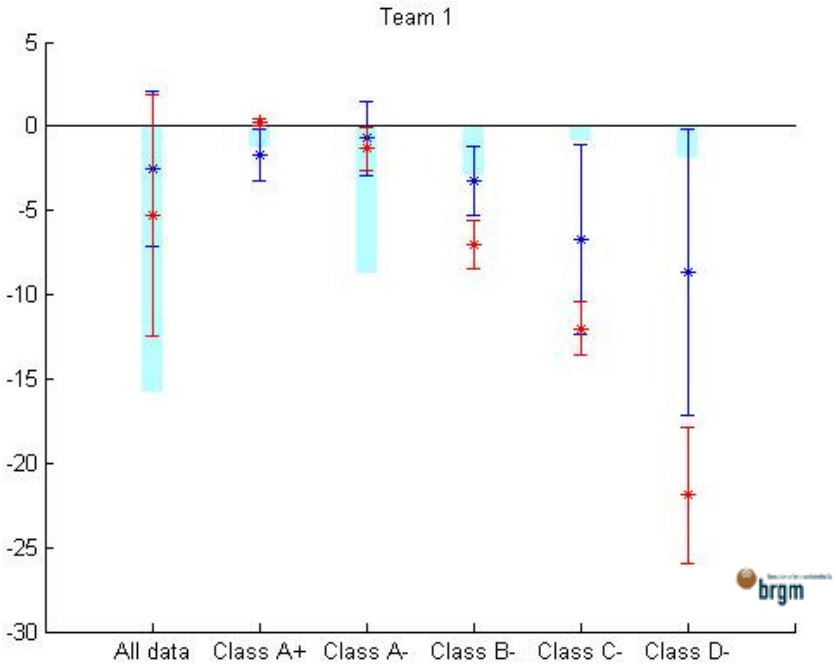


Figure 4.4.12: Distribution of velocities estimated from PS (blue) and Levelling data (red) within each class. The cyan bars show the number of levelling points contained in each class. This graph allows identifying disagreements of estimated PS velocities with respect to the velocities estimated from levelling measurements. Y axis is mm/y.

Team 2

Class	Mean velocities levelling (mm/year)	Standard deviation velocities levelling (mm/year)	Mean velocities PS (mm/year)	Standard deviation velocities PS (mm/year)	Number Points
All	-6.8	8.2	-2.4	3.3	478
A +	0.5	0.7	-0.5	1.2	33
A -	-1.5	1.4	-1.0	1.9	238
B -	-7.3	1.5	-3.1	2.8	87
C -	-11.8	1.3	-5.8	4.2	41
D -	-22.7	5.2	-5.0	4.0	79

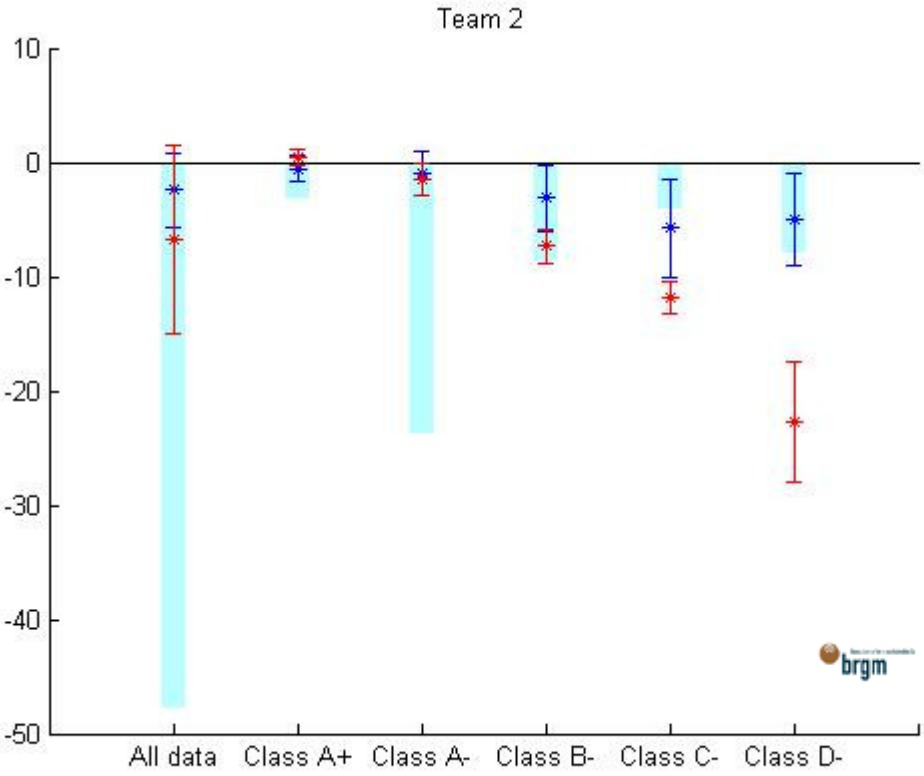


Figure 4.4.13. Distribution of velocities estimated from PS (blue) and Levelling data (red) within each class. The cyan bars show the number of levelling points within in each class. This graph allows identifying disagreements of estimated PS velocities with respect to the velocities estimated from levelling measurements. Y axis is mm/y.

Team 3

Class	Mean velocities levelling (mm/year)	Standard deviation velocities levelling (mm/year)	Mean velocities PS (mm/year)	Standard deviation velocities PS (mm/year)	Number Points
All	-5.0	7.0	-1.1	1.7	328
A +	0.7	1.1	-0.3	0.9	35
A -	-1.6	1.4	-0.6	1.0	186
B -	-7.2	1.6	-1.8	1.6	52
C -	-11.7	1.5	-3.6	3.7	21
D -	-22.5	4.6	-2.0	2.0	34

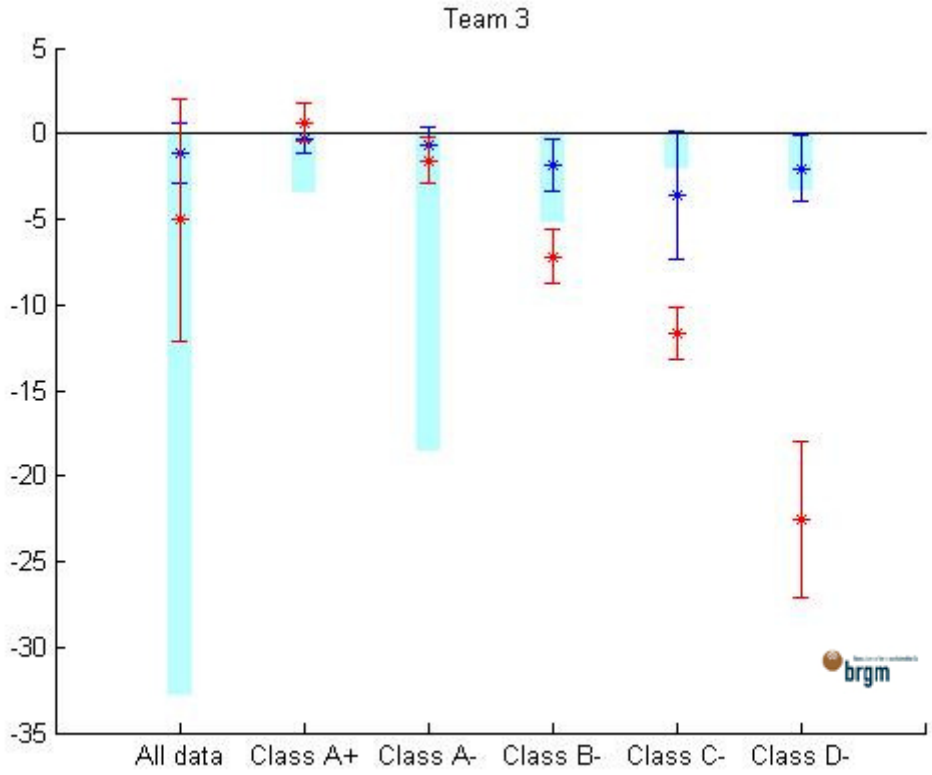


Figure 4.4.14: Distribution of velocities estimated from PS (blue) and Levelling data (red) within each class. The cyan bars show the number of levelling points within in each class. This graph allows identifying disagreements of estimated PS velocities with respect to the velocities estimated from levelling measurements. Y axis is mm/y.

Class	Mean velocities levelling (mm/year)	Standard deviation velocities levelling (mm/year)	Mean velocities PS (mm/year)	Standard deviation velocities PS (mm/year)	Number Points
All	-5.4	7.3	-1.7	2.8	447
A +	0.6	1.0	-0.5	1.4	42
A -	-1.5	1.4	-0.6	1.6	248
B -	-7.2	1.5	-2.6	2.2	69
C -	-12.0	1.4	-4.7	4.2	37
D -	-22.2	5.8	-4.3	3.8	51

1. Team 4

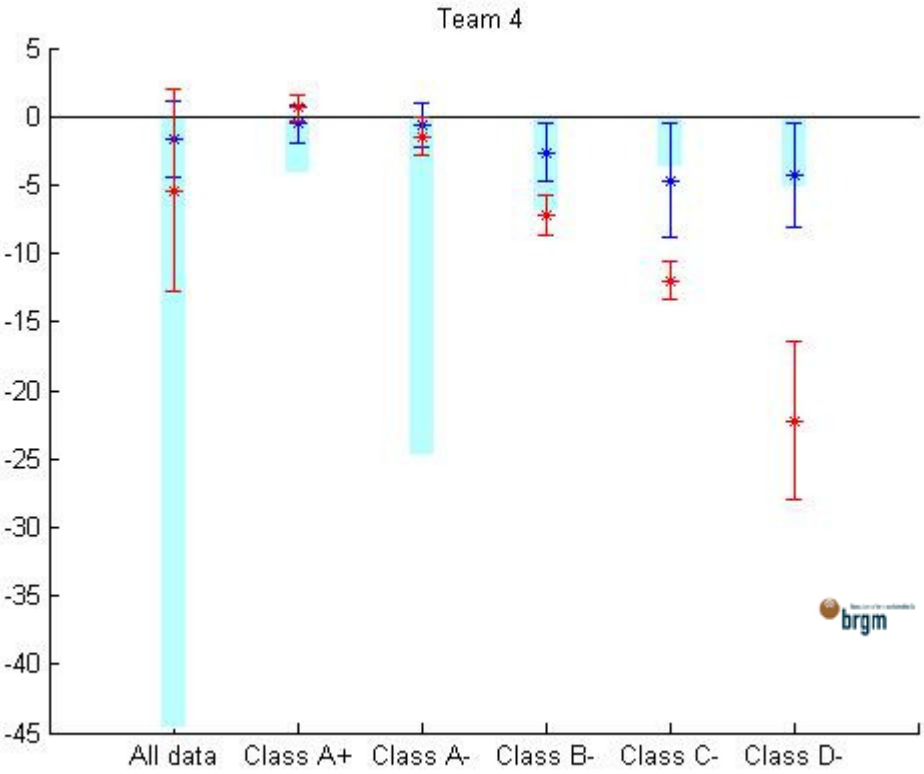


Figure 4.4.15: Distribution of velocities estimated from PS (blue) and Levelling data (red) within each class. The cyan bars show the number of levelling points within in each class. This graph allows identifying disagreements of estimated PS velocities with respect to the velocities estimated from levelling measurements. Y axis is mm/y.

2. Team 5

Class	Mean velocities levelling (mm/year)	Standard deviation velocities levelling (mm/year)	Mean velocities PS (mm/year)	Standard deviation velocities PS (mm/year)	Number Points
All	-6.5	7.8	-3.0	2.5	348
A +	0.3	0.5	-1.4	1.3	29
A -	-1.6	1.4	-1.8	1.7	174
B -	-7.1	1.6	-3.6	1.8	59
C -	-11.9	1.5	-6.2	3.4	29
D -	-21.7	4.4	-5.0	2.3	57

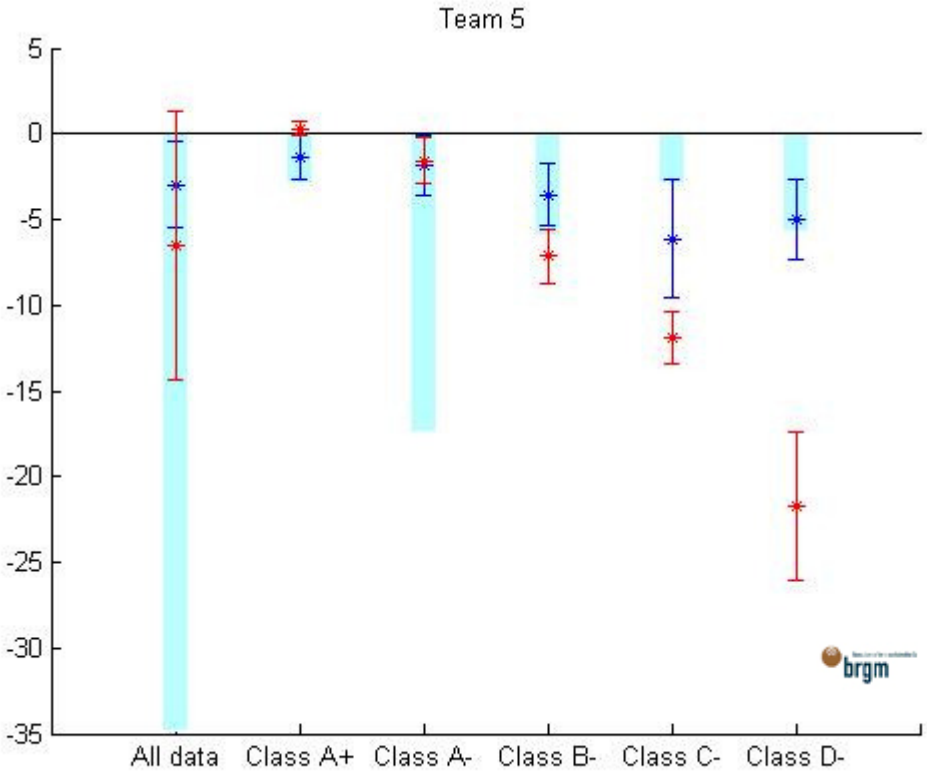


Figure 4.4.16: Distribution of velocities estimated from PS (blue) and Levelling data (red) within each class. The cyan bars show the number of levelling points within in each class. This graph allows identifying disagreements of estimated PS velocities with respect to the velocities estimated from levelling measurements. Y axis is mm/y.

Team 6

Class	Mean velocities levelling (mm/year)	Standard deviation velocities levelling (mm/year)	Mean velocities PS (mm/year)	Standard deviation velocities PS (mm/year)	Number Points
All	-3.9	5.5	-0.7	0.7	136
A +	0.3	0.5	-0.4	0.4	18
A -	-1.4	1.2	-0.6	0.7	77
B -	-6.7	1.4	-0.6	0.5	24
C -	-11.7	1.8	-1.6	0.6	8
D -	-19.6	4.3	-1.6	0.5	9

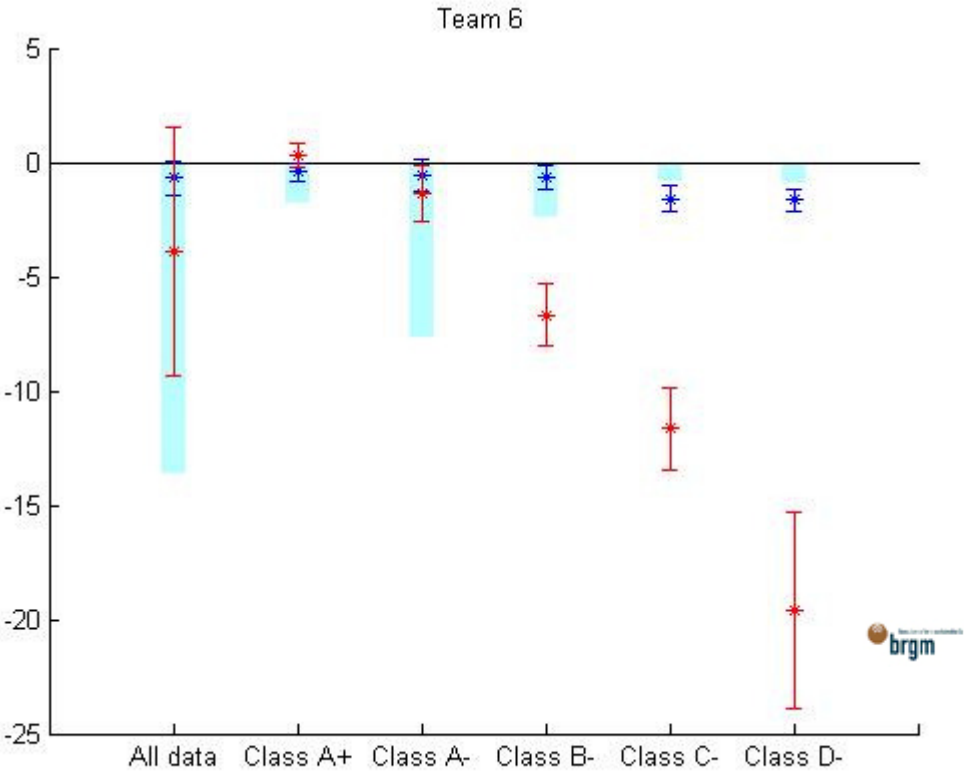


Figure 4.4.17: Distribution of velocities estimated from PS (blue) and Levelling data (red) within each class. The cyan bars show the number of levelling points within in each class. This graph allows identifying disagreements of estimated PS velocities with respect to the velocities estimated from levelling measurements. Y axis is mm/y.

Team 7

Class	Mean velocities levelling (mm/year)	Standard deviation velocities levelling (mm/year)	Mean velocities PS (mm/year)	Standard deviation velocities PS (mm/year)	Number Points
All	-5.4	7.2	-2.8	3.2	417
A +	0.6	1.0	-0.6	1.2	38
A -	-1.4	1.3	-1.5	1.8	221
B -	-7.1	1.5	-3.6	2.2	75
C -	-11.7	1.3	-6.6	3.9	37
D -	-22.0	5.5	-6.8	3.9	46

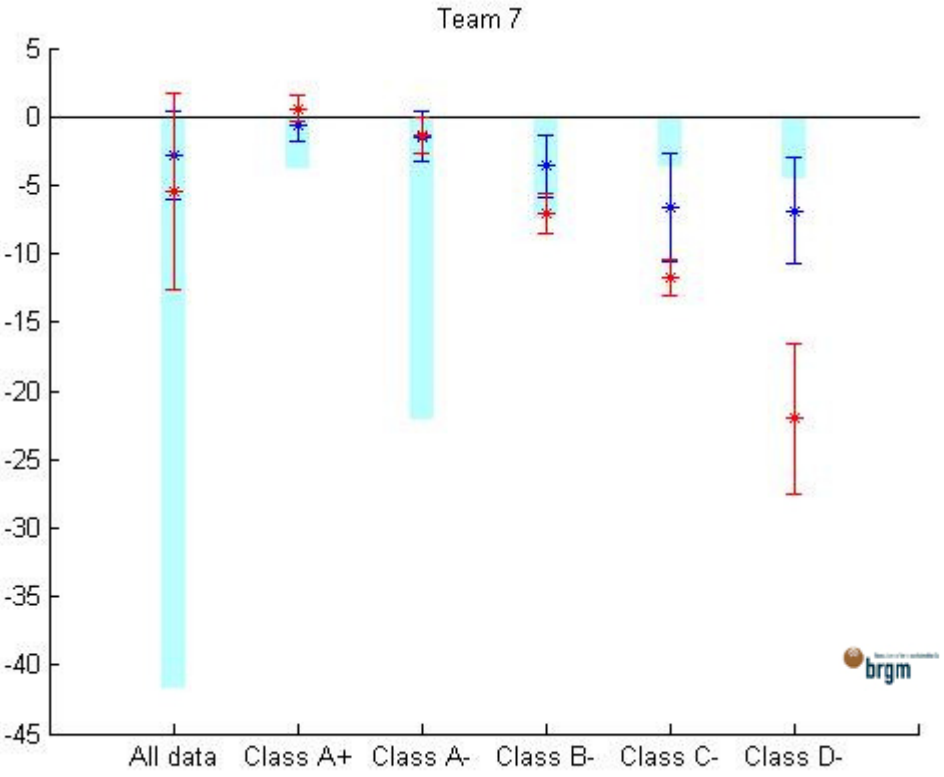


Figure 4.4.18: Distribution of velocities estimated from PS (blue) and Levelling data (red) within each class. The cyan bars show the number of levelling points within in each class. This graph allows identifying disagreements of estimated PS velocities with respect to the velocities estimated from levelling measurements. Y axis is mm/y.

Team 8

Class	Mean velocities levelling (mm/year)	Standard deviation velocities levelling (mm/year)	Mean velocities PS (mm/year)	Standard deviation velocities PS (mm/year)	Number Points
All	-9.2	10.0	-4.4	5.7	817
A +	0.6	0.9	0.0	2.0	51
A -	-1.5	1.4	-0.5	3.1	334
B -	-7.2	1.5	-4.0	3.5	151
C -	-11.9	1.4	-7.3	4.8	73
D -	-24.4	5.9	-11.3	3.7	208

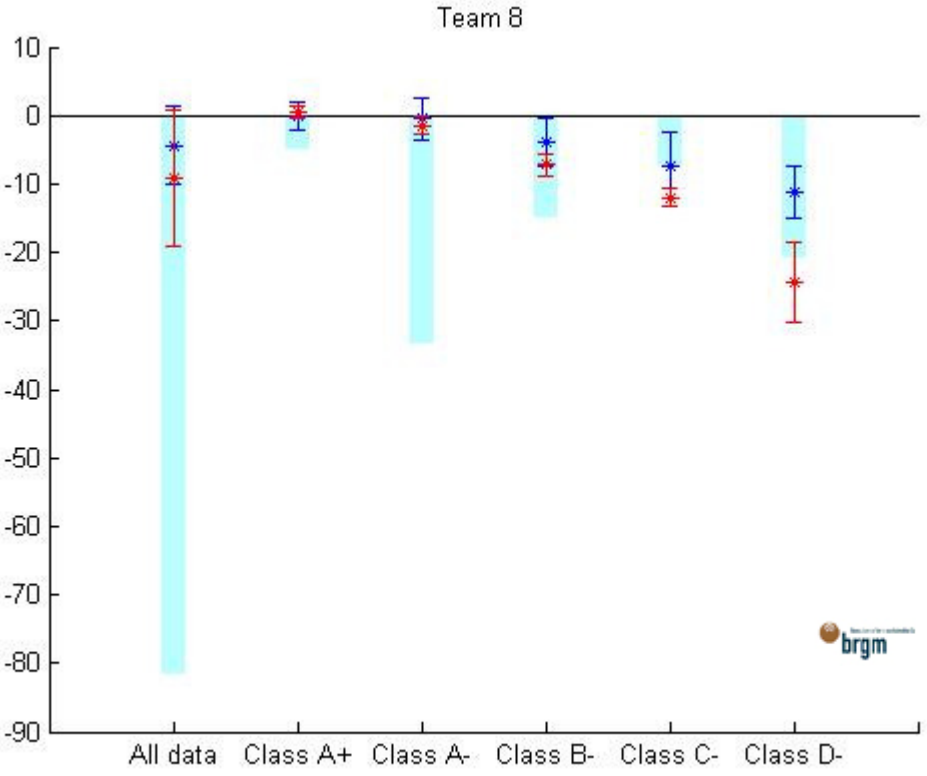


Figure 4.4.19: Distribution of velocities estimated from PS (blue) and Levelling data (red) within each class. The cyan bars show the number of levelling points within in each class. This graph allows identifying disagreements of estimated PS velocities with respect to the velocities estimated from levelling measurements. Y axis is mm/y.

Statistics on PS and levelling velocity differences

The table below summarises the overall differences between PS and levelling velocities. Differences are expressed as mean and standard deviation of $T_{(i)}$ minus levelling, for each team and for all points.

Team	Number of points	Mean|Ti – Levelling|	STDV Ti - Levelling
T1	158	2.8	5.3
T2	478	4.4	7.2
T3	328	3.9	6.6
T4	447	3.7	6.3
T5	348	3.6	6.8
T6	136	3.2	5.1
T7	417	2.6	5.7
T8	817	4.7	6.7

Table 4.4.IX : overall differences between PS and levelling velocities

The figures below offer a graphical overview of the overall accuracy of PS velocity measurements against the levelling velocity measurements. The mean of the difference between PS velocity and levelling velocity is plotted against the number of analysed PS for each team. The trend shows that PS values belonging to teams that provided a larger coverage of PS on the deforming area have a lower overall accuracy. That illustrates a trade-off between density and accuracy of the measure. The graph below should be interpreted by taking into account the results of the next paragraph.

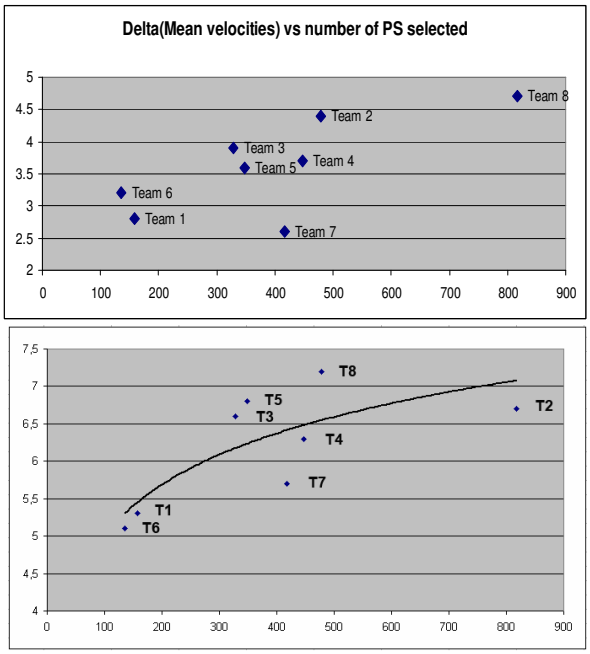


Table 4.4.X. **Y axis** = mean (above) and standard deviation (bottom) of Vps – Vlev ; **X axis** = number of PS interpolated at levelling location

4.5 Overlying with ancillary data

There was no clear relationship between the geological faults in the area and the subsidence maxima as shown in the Diffsar or PSinsar results. The most telling data are the Diffsar measurements since this formed a continuous dataset. However it is possible to see on Figure 4.5.1 that the subsidence forms a north to south band, which has been sinistrally offset by a curvilinear east-west trending fault. This offset might be explained as the expected result of a normal geological fault which downthrows to the south, and which cross-cuts the moderately WNW dipping rock strata. However it is also clear that this maximum subsidence is also related to the westernmost mined panels of both the northern and southern areas.

It was important to first gain an overall impression of which mining panels were likely to be affecting the ground motion as shown by the PS data. To this end we overlaid the panels on the Diffsar data and gave the panels a number to make referencing easier.

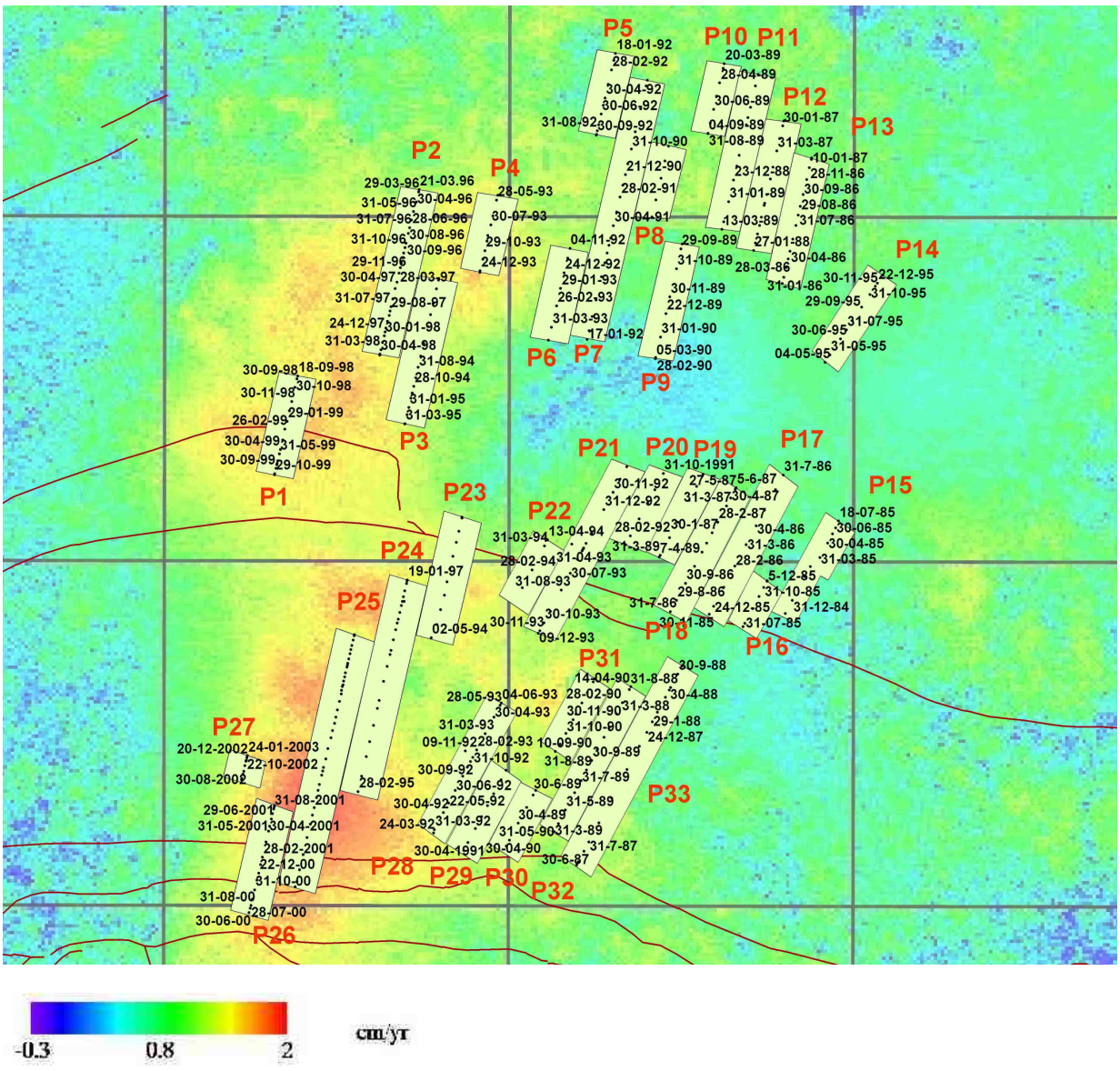


Figure 4.5.1. Mining panels, with dates, overlaid on the Diffsar data. Red lines are geological faults.

From the above figure it is apparent that panels 1-4 and 24-27 are most likely to be affected by subsidence. It was then necessary to look at he PS point distributions to see which panels coincided with PS points from each team's data. Panel one was found to be overlain by points for Teams 1, 2, 3 4, 7 and 8. These points were extracted and their motion histories plotted. We could then examine any relationships between dates of mining and ground motions detected by the PSinSAR analysis.

Panel one started to be mined on the 18/9/98, mining finished on the 18/11/99. The following graphs show the point histories for all the points, of a particular team, found to coincide with panel 1.

Team One

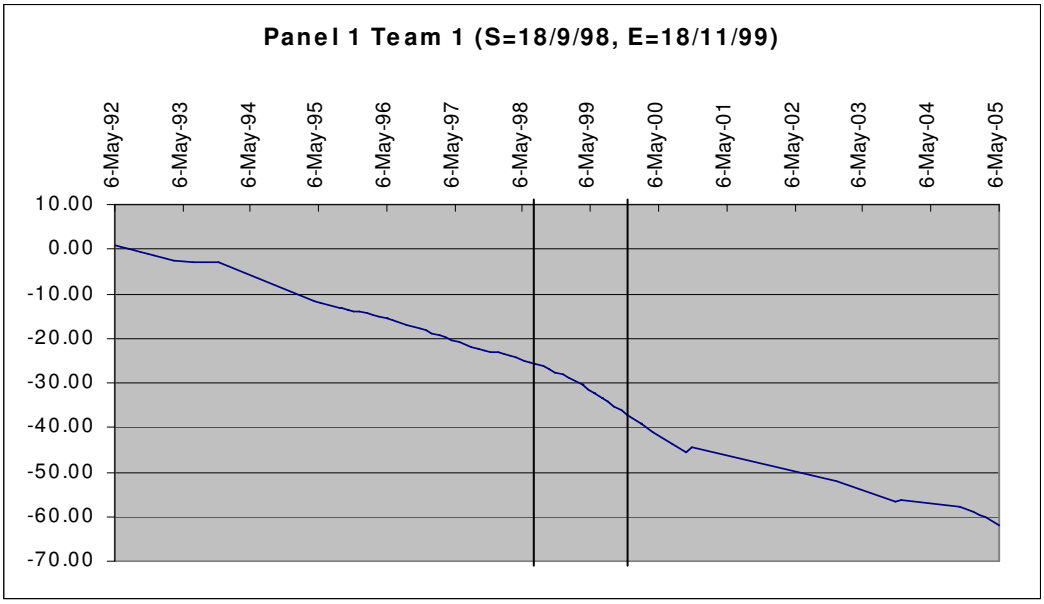


Figure 4.5.2 Graph showing PS Point history for Team 1 above Mining Panel 1

The two vertical lines show the start and end of mining of this panel. Point histories show fairly constant linear ground motion before and after this panel had been mined. It is possible that the ground motion shown by this point is a cumulative result of the earlier mining of other nearby panels. Panel 6, which is approximately 1800m to the northeast (Figure 4.5.1), was mined in 1992 and 1993 and panels 2, 3 and 4, which lie between panel 1 and panel 6 were mined in 1993-1998. It would therefore seem, if the PS results are accurate, that the surface effect of mining is not only restricted to immediately above a mined panel.

Team Two

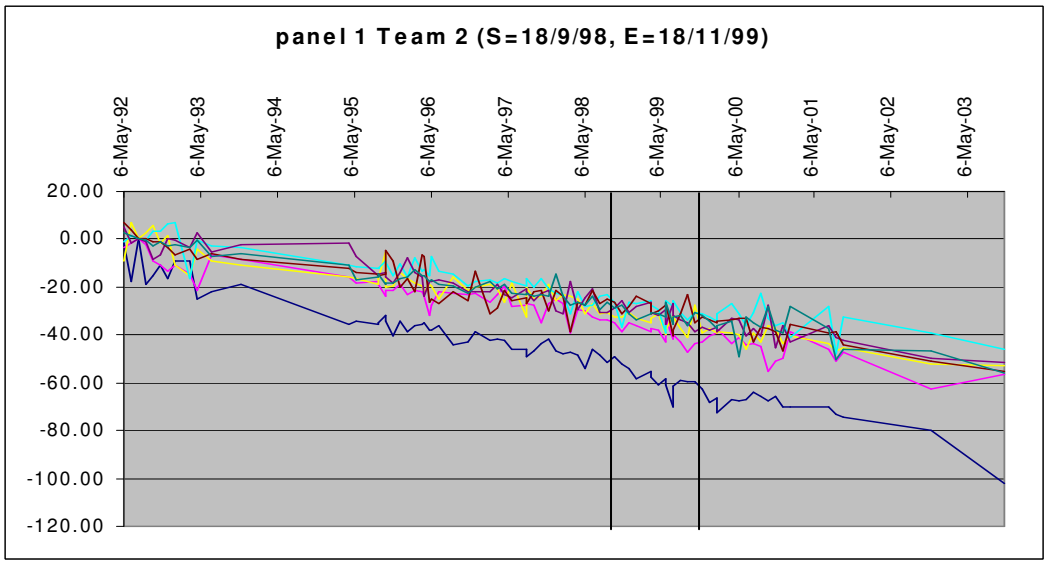


Figure 4.5.3 Graph showing PS Point histories for Team 2 above Mining Panel 1

The PS points found over panel 1 by Team Two all have a similar trend of subsidence. Once again subsidence has started a few years before the panel was mined, this is possibly combined effect of the earlier exploitation of nearby panels.

Team Three

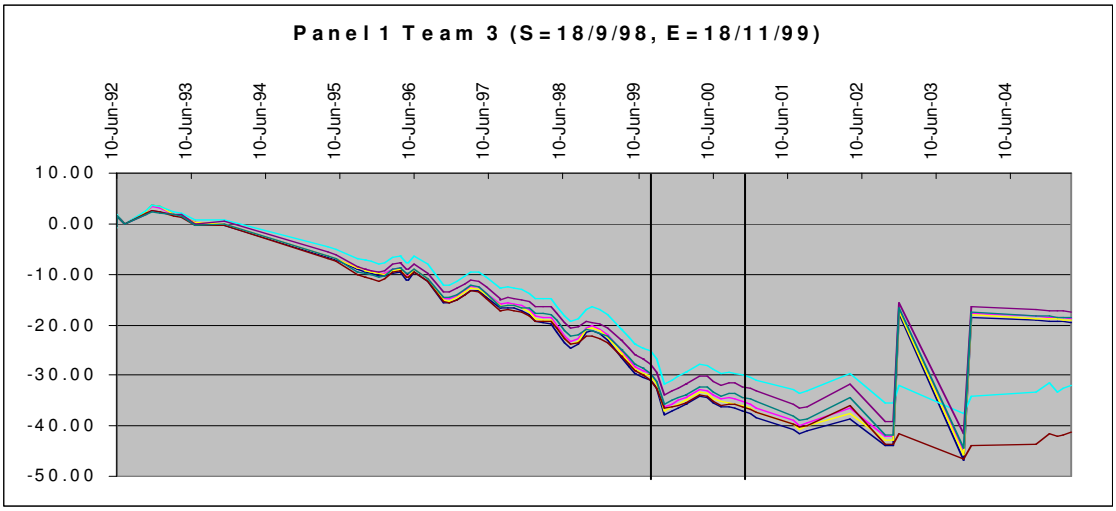


Figure 4.5.4 Graph showing PS Point histories for Team 3 above Mining Panel 1

The PS points found for the panel 1 area by Team 3 also show the apparent influence of nearby earlier mining. However in this teams PS dataset there is a steepening of the line shown in the graph after the start of mining and then a flattening of the line once the panel mining has been completed. This might represent the direct influence of the mining on the ground above it, with faster subsidence during the period the panel was mined and then a

drop of in the subsidence rate once the panel was completed. Some strange results towards the end of the period processed are present.

Team Four

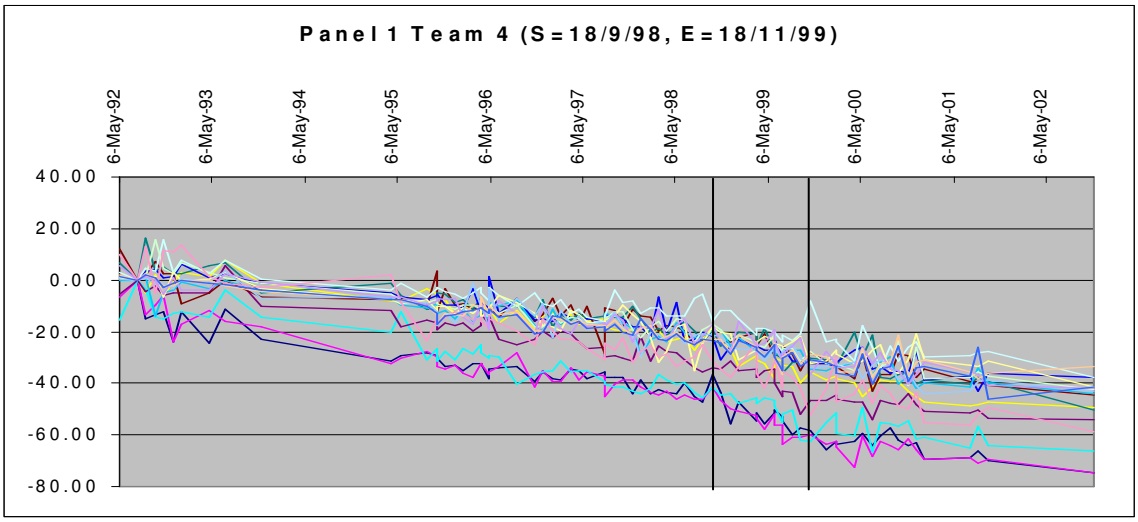


Figure 4.5.5 Graph showing PS Point histories for Team 4 above Mining Panel 1

Team Four's points show the same subsidence trend as Team Three's data. We can see an overall linear subsidence before the mining of panel one commenced. Once mining commenced we can see an increase in subsidence rate followed by a flattening off of the rate once mining had stopped. This kind of subsidence behaviour can fit with the ground conditions. Although all the PS points found above panel one show a similar trend of subsidence there is a large variation of actual values.

Team Five

No PS points over Panel One.

Team Six

No PS points over Panel One.

Team Seven

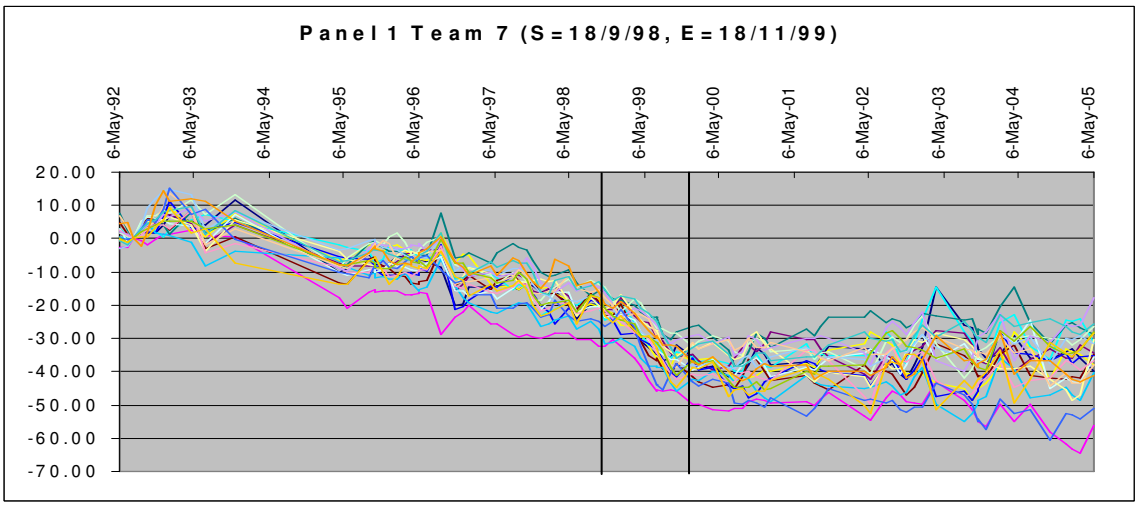


Figure 4.5.6 Graph showing PS Point histories for Team 7 above Mining Panel 1

Team Seven's points show the same subsidence trend as Team Three's data. We can see an overall linear subsidence before the mining of panel one commenced. Once mining commenced we can see an increase in subsidence rate followed by a flattening off of the rate once mining had stopped. After the end of 1999 the subsidence history of the points shows an increasing amount of variation. Team Seven also show a fairly good match with mining dates for Panel 3, although there is still subsidence before and after the exploitation of the panel. Team Seven also show a fairly good match with mining dates for Panel 3, although there is still subsidence before and after the exploitation of the panel.

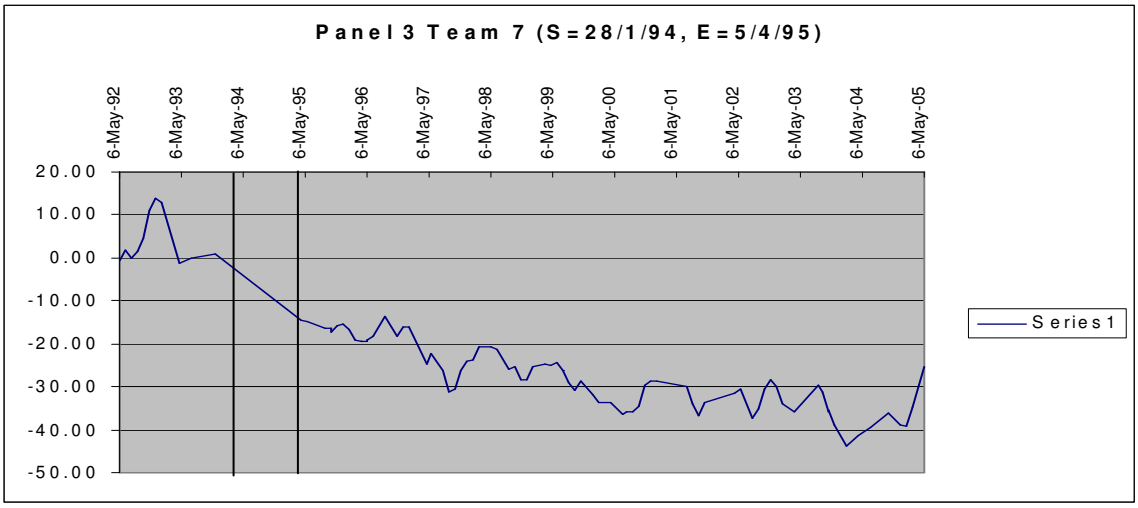


Figure 4.5.7 Graph showing PS Point histories for Team 7 above Mining Panel 3

Team Seven also shows a fairly good match with mining dates for Panel 24

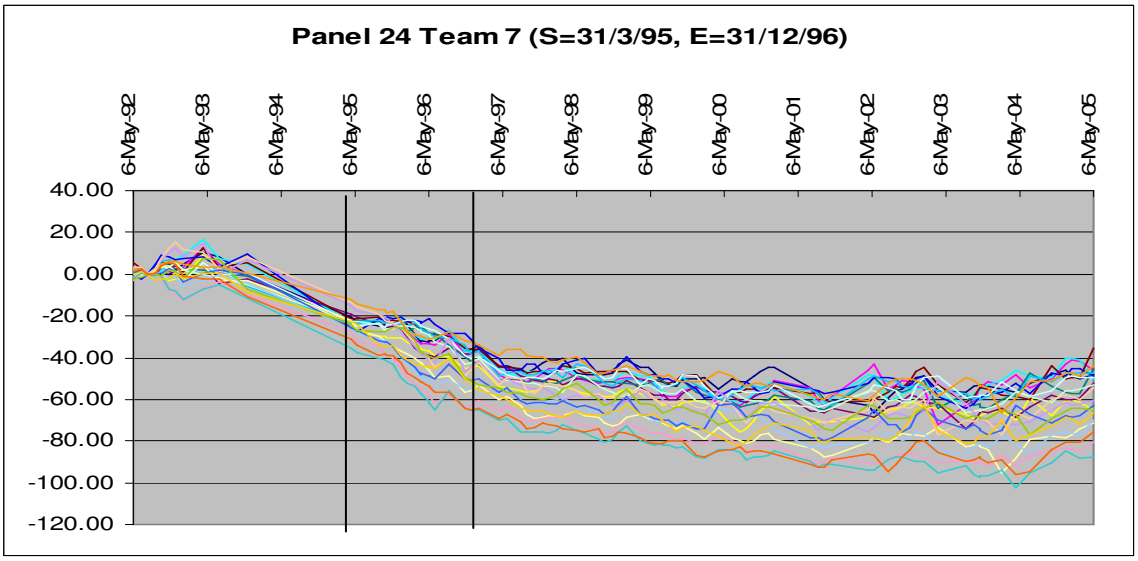


Figure 4.5.8 Graph showing PS Point histories for Team 7 above Mining Panel 24

Team Eight

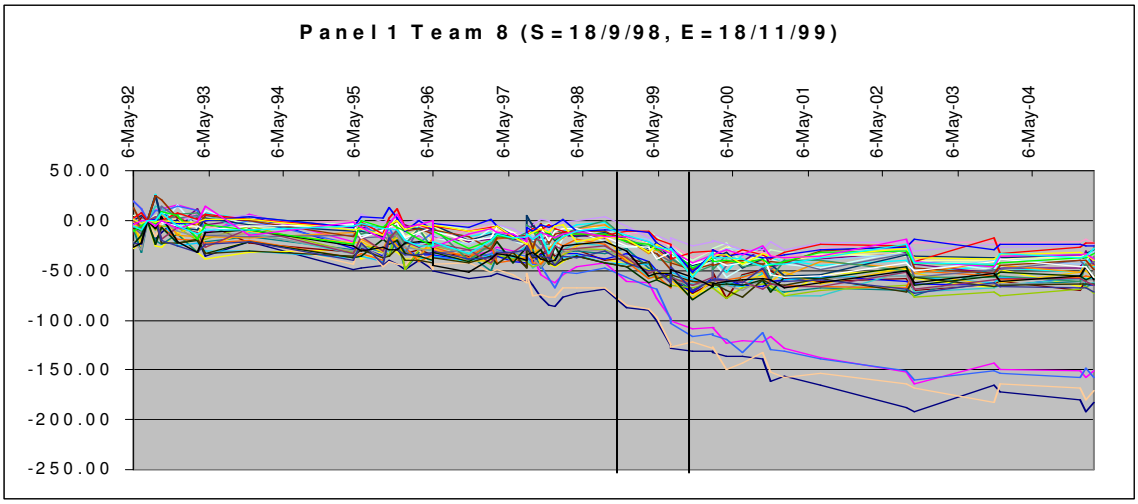


Figure 4.5.9 Graph showing PS Point histories for Team 8 above Mining Panel 1

Team eight's PS histories also show a general linear subsidence that starts befote the panel was mined. There is a marked increase in subsidence rate when mining began and the subsidence rate slows or even stops when mining ceases. There are 4 points that show the subsidence carried on for an extra 100mm. The reason for this is not known.

Some of Team Eight's PS histories also show a general linear subsidence that starts before the panel was mined. There is a marked the point histories show good correspondence to the dates of mining. This example below is Team Eight's data for Panel 2;

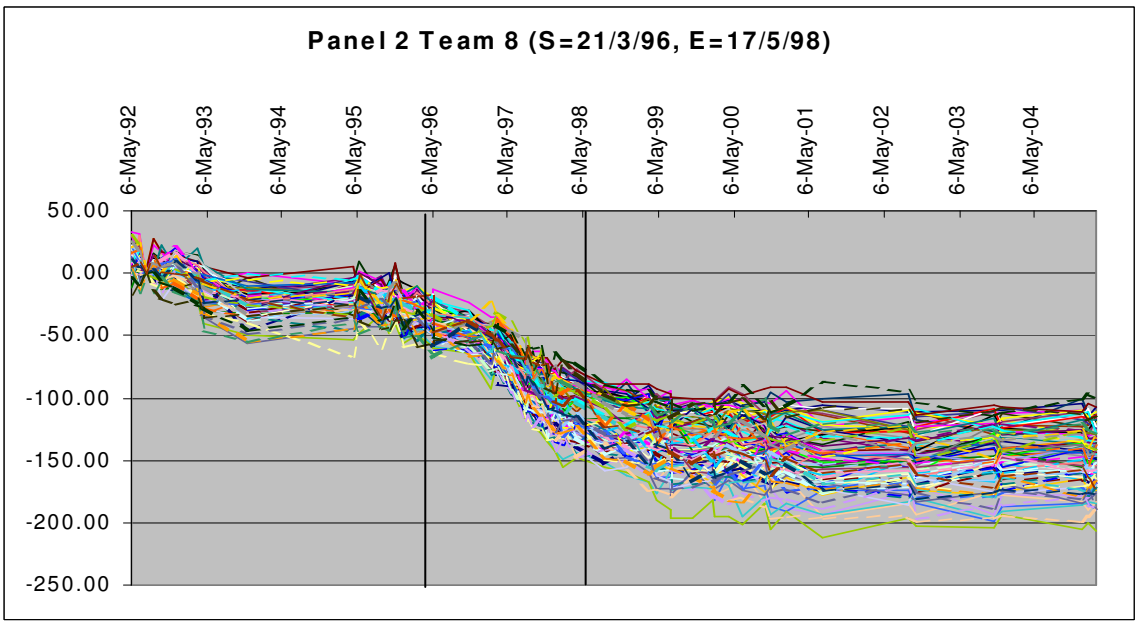


Figure 4.5.10 Graph showing PS Point histories for Team 8 above Mining Panel 2

This data for Team Eight over Panel 2 shows a small subsidence before mining started in this panel (prior to March 1996), this small amount can be from neighbouring earlier mining of panels 3, 4 and 5 (Figure 3.4.1). After mining starts in March 1996 there is a marked increase in subsidence, which is fairly linear for all points until the end of mining in May 1998. After May 1998 the subsidence slows and then levels out.

However Team Eight do not always have such good results, the figure below are Team Eight's results from Panel 3:

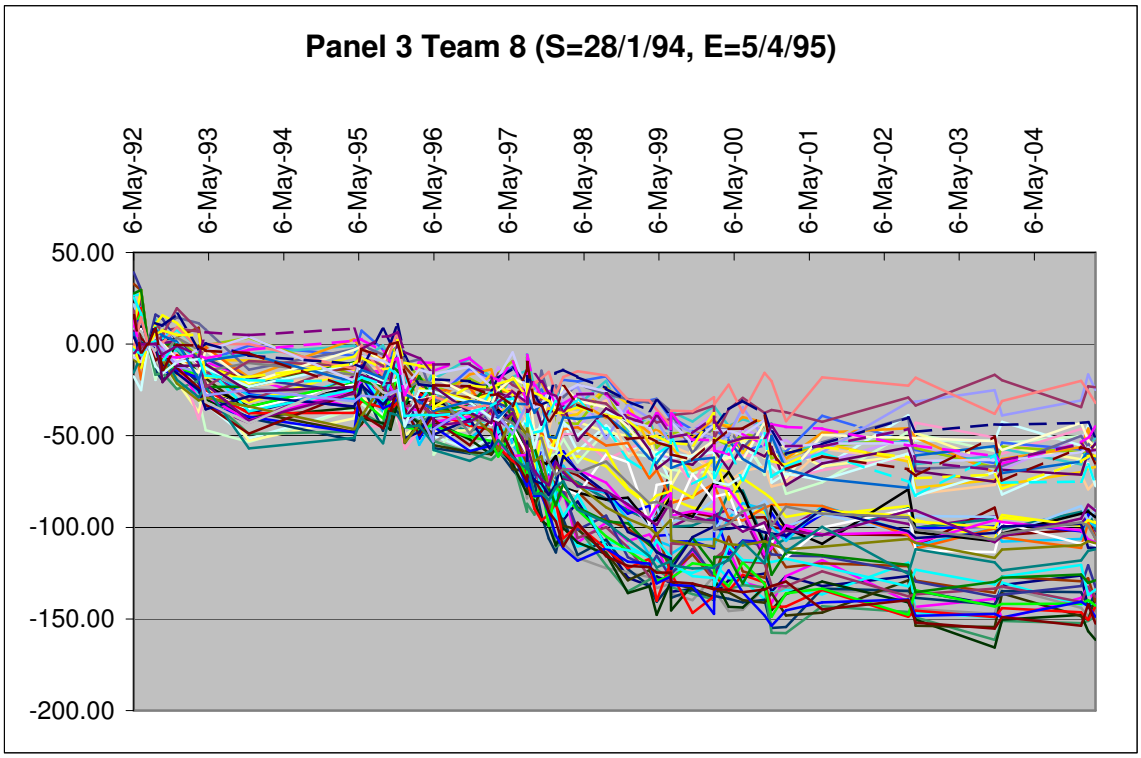


Figure 4.5.11 Graph showing PS Point histories for Team 8 above Mining Panel 2

Team Eight results over this panel show no subsidence at all durng the period of mining. The reason for this is that no SAR scenes were available during this time so that they were not processed. Due to the nature of the processing, large gaps in SAR availability causes problems such as this.

Conclusions for overlying with ancillary data

All the teams who have permanent scatterer points within the areas of the mined panels show subsidence. However the subsidence nearly always started before the mining for that panel took place. This subsidence that occurred before the area was directly undermined, does seem to relate in some way to previous mining activities, which has occurred in nearby

panels (within 1.8km). Though we have no information on the processing, we can wonder if it is not an effect of a linearization of the deformation *i.e.* the phase temporal profile would have been adjusted to better fit to an ideal model which do not correspond with the actual less regular deformation.

5. Intercomparison activities

The intercomparison activities are meant to compare PS data values provided by a processing team with each other, as a relative assessment. The goal of this exercise is to provide an estimate of how well the 8 Teams agree with each other. Ideally, they are supposed to provide - using the same SAR data set – very similar results. Since in this exercise there is no direct comparison with *ground truth* (i.e. levelling network), no data interpolation has been performed as we compare the original PS datasets provided by each team.

We compare the following quantities;

- Spatial distribution and densities
- Atmospheric Phase Screens (APS)
- Geocoding of data
- Line of sight velocities

5.1 PSI spatial distribution and density

Output:

- Plots of overall PS density for each team
- Density value (Number of PS per square kilometre) overall, and for rural, urban and unstable areas
-Visual plots of density in urban and unstable areas

The area processed by each team was examined. The convex hulls of the PS points produced for each team are shown in Figure 5.1.1 below.

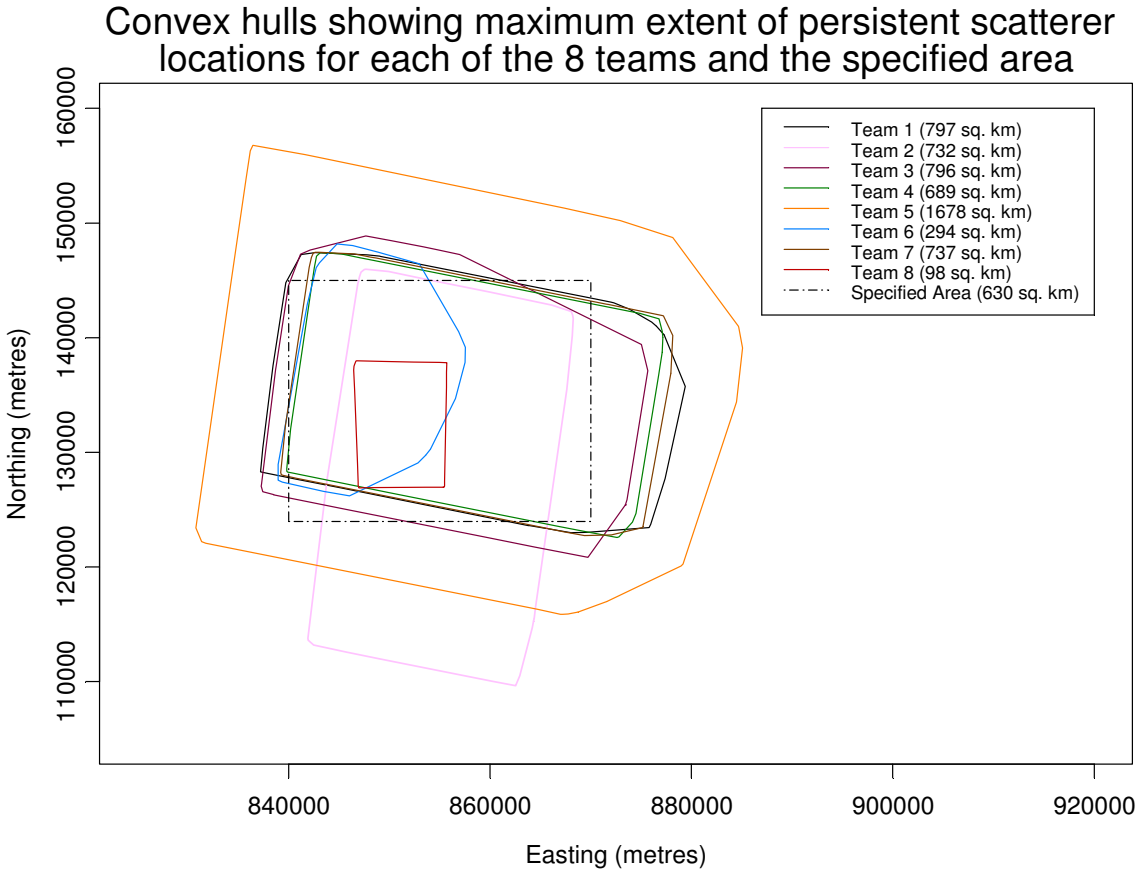


Figure 5.1.1 Areas processed by each team

It can be seen in this figure that Team 5 processed the largest area and Team 8 the smallest. Also worth noting is the sloping top and bottom edges of the processed areas. These slopes correspond to the SAR geometry.

To derive the PS density for each team it was first necessary to identify an area in which all teams had PS measurements. Since Team 8 and 6 had processed the smallest area it was their results, which controlled the selection of an area to make the density measurements in. The Task 6 report details the methodology.

Once a measurement area was selected its size was measured and found to be 33.87 km squared. A GIS system was then used to count the points from each team that fell into the area. The density of PS points within this measurement area could then be calculated.

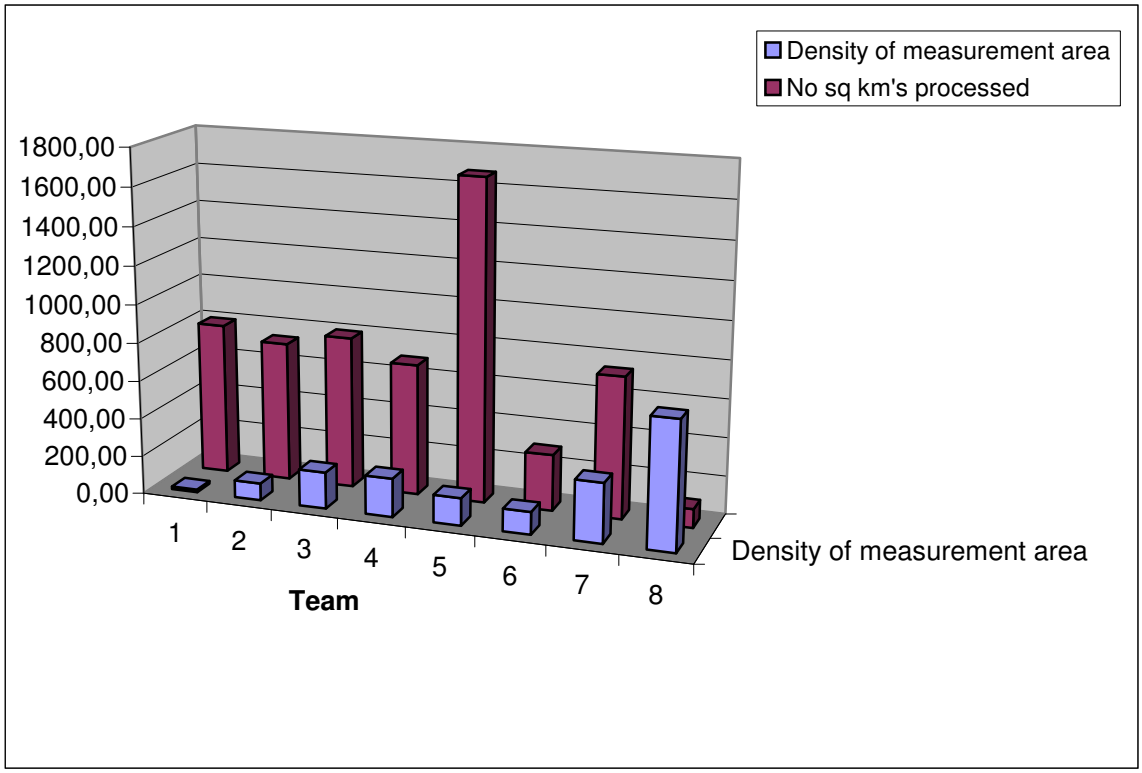


Figure 5.1.2. Graphical representation of the size of area processed by each team and the density of PS points produced by the teams. Units is PS/km².

Team 1 produced the lowest overall density of PS points, while Team 8 produced the highest density. However, while it is true that Team 8 processed the smallest area and had the highest density, it does not necessarily follow that processing smaller areas produces higher PS densities. Team One, with the lowest density did not process the smallest area. It must be remembered that both Team 1 and Team 8 seem to have used different processing strategies than the other teams (this assumption is based on the differences in the spatial patterns of the PS; Team 1's PS seem to be on a grid and Team 8 have many PS which cannot be related to ground features). Team 6 having a lower density than Team 7 supports the fact that there is no relationship between area processed and density.

In order to test the distribution of these densities with respect to land use it was necessary to digitise the urban areas within the defined measurement area rectangle. The size of the urban areas was found to be 11.69km squared, and be subtracting this from the size of the measurement area the rural areas were found to be 22.18km squared. Once points falling within the urban areas were extracted and counted it was possible to derive a density value of PS points in rural and urban areas for each team (Table 4.1.1).

For each team the highest densities were always seen in the urban areas. This was expected since this is where most features likely to act as permanent scatterers are to be found.

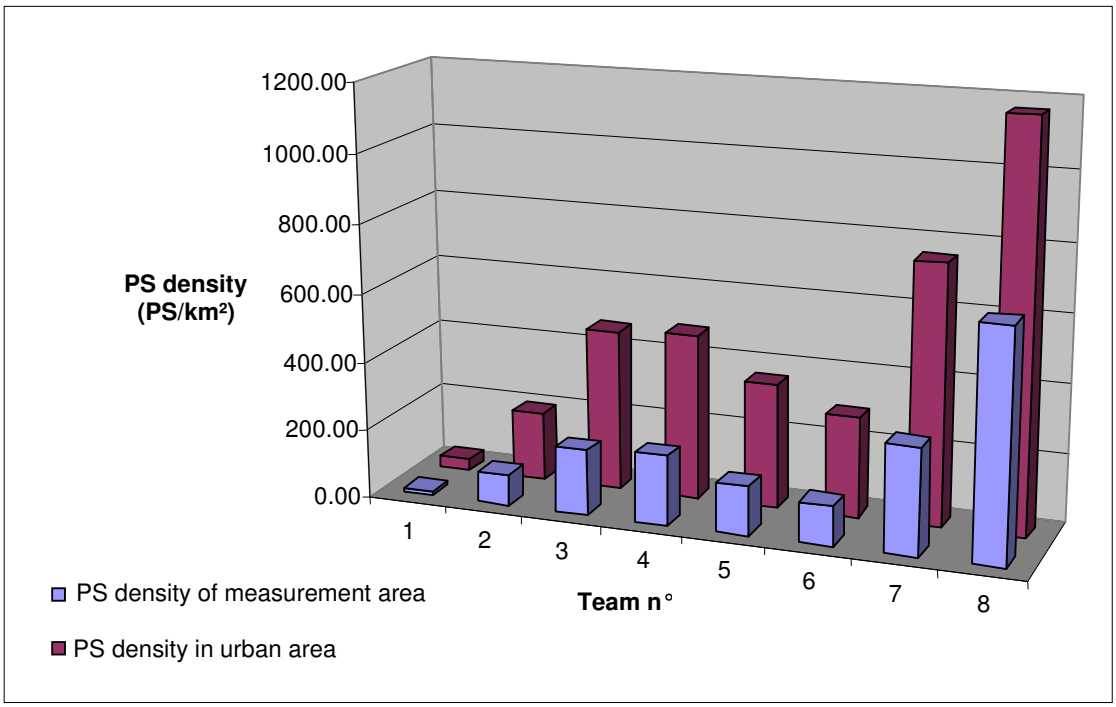


Figure 5.1.3. Density of PS points in the overall measurement area and that of urban areas.

It is also apparent that high overall densities also mean high urban densities.

The density measurement area was then divided into 100m square cells so that we could produce a colour-coded density for each team. Colour coding was applied as a colour stretch across a team's dataset, therefore colours are not directly comparable between teams but it does highlight which areas have a high density within a dataset. DiffSAR results were used to define areas of rapid movement and these were overlaid onto the colour coded densities.

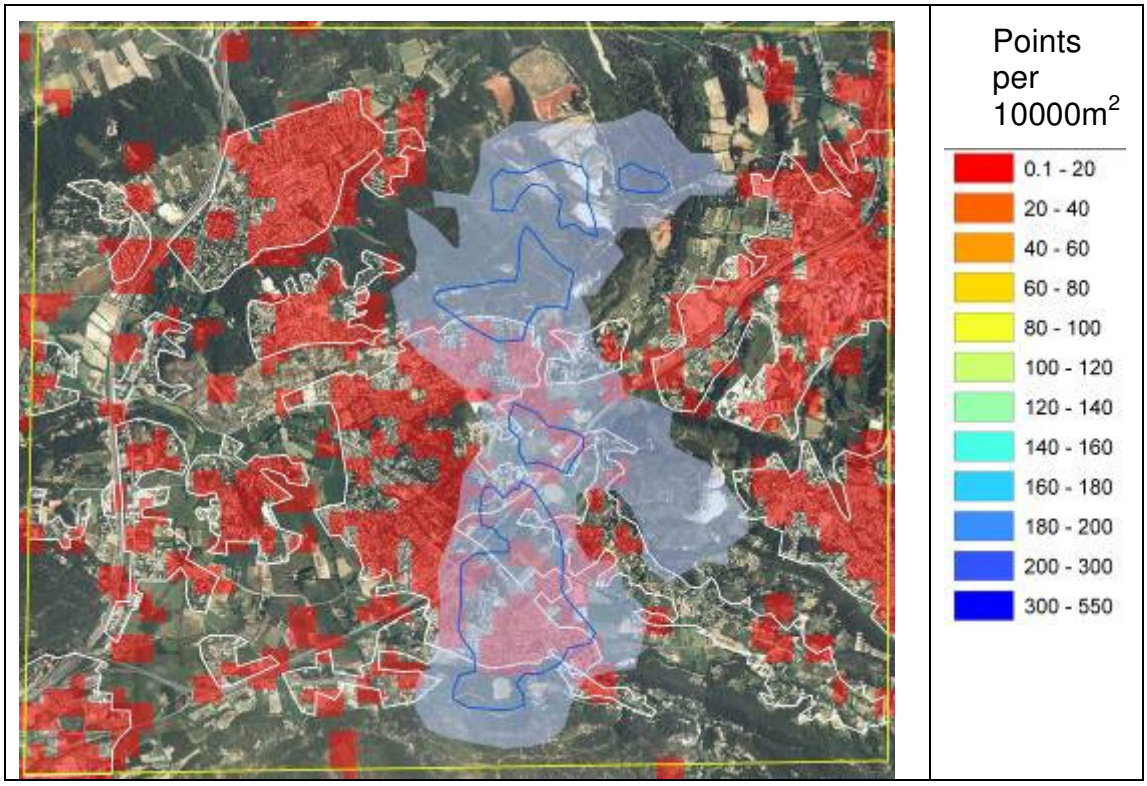


Figure 5.1.4 Team 1 density. 10000m2 density; blue areas are high PS density within the team 1 dataset, red area lower density areas. Clear areas have no PS points. The yellow rectangle is the area defined for the overall density measurements. White polygons are urban areas, blue and red polygons are areas of subsidence, with red areas moving more rapidly than the blue areas.

Figure 5.1.4 shows that Team 1 has a higher density of PS in the urban areas than the rural areas. This is true for all teams. Interestingly it also shows a relatively high density of PS points in the southern red area. This is an area of rapid subsidence and Team 1 were able to detect PS points here. Red areas to the north do not have any points present, however this is to be expected since this is a rural area, therefore the number of features that could act as permanent scatterers will be limited. It must be remembered that Team One's PS points are presented on a grid bases, this accounts for the seemingly large amount of blue 'high density' areas.

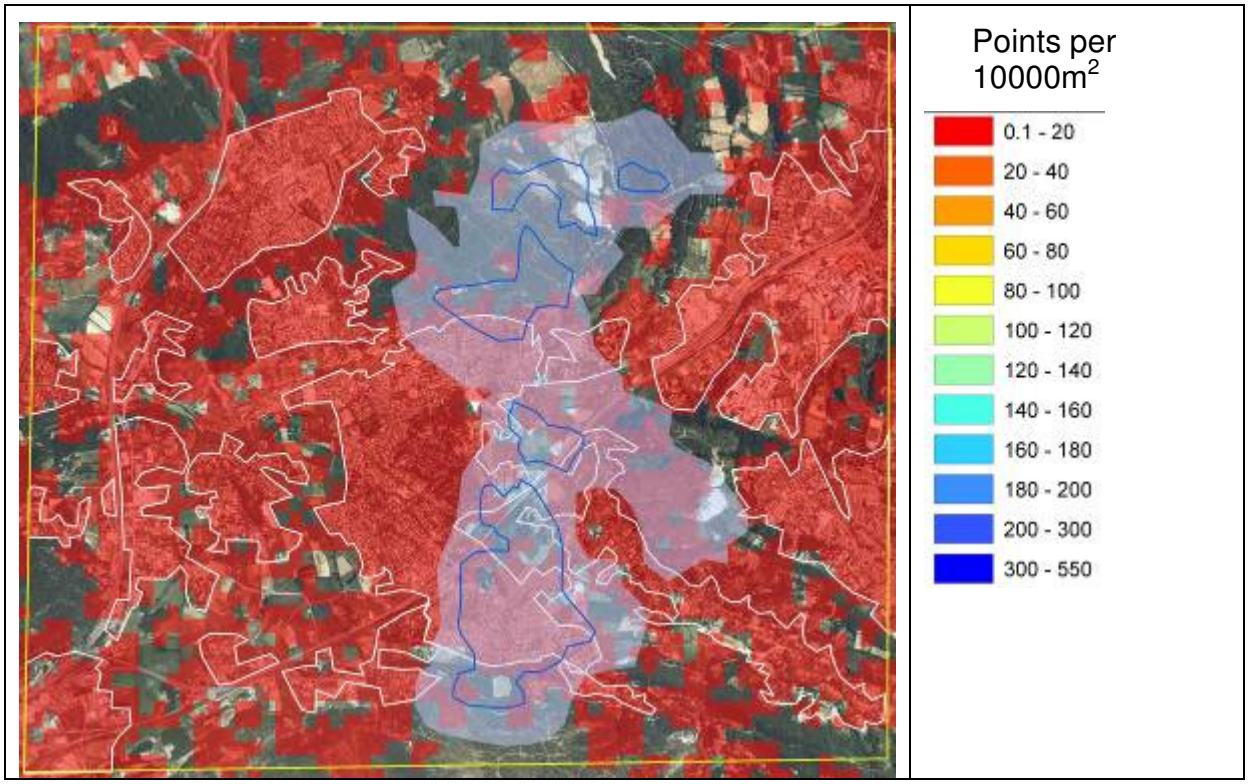


Figure 5.1.5 Team 2 densities. 10000m2 density; blue areas are high PS density within the team 2 dataset; red are lower density areas. Clear areas have no PS points. The yellow rectangle is the area defined for the overall density measurements. White polygons are urban areas, blue and red polygons are areas of subsidence, with red areas moving more rapidly that the blue areas.

Team Two's results show a higher number of points in rural areas than Team One. Once again the highest density is in the centre of the urban areas. There are still a number of points in the areas of most rapid subsidence, so Team Two was able to detect some PS pints within areas of rapid subsidence.

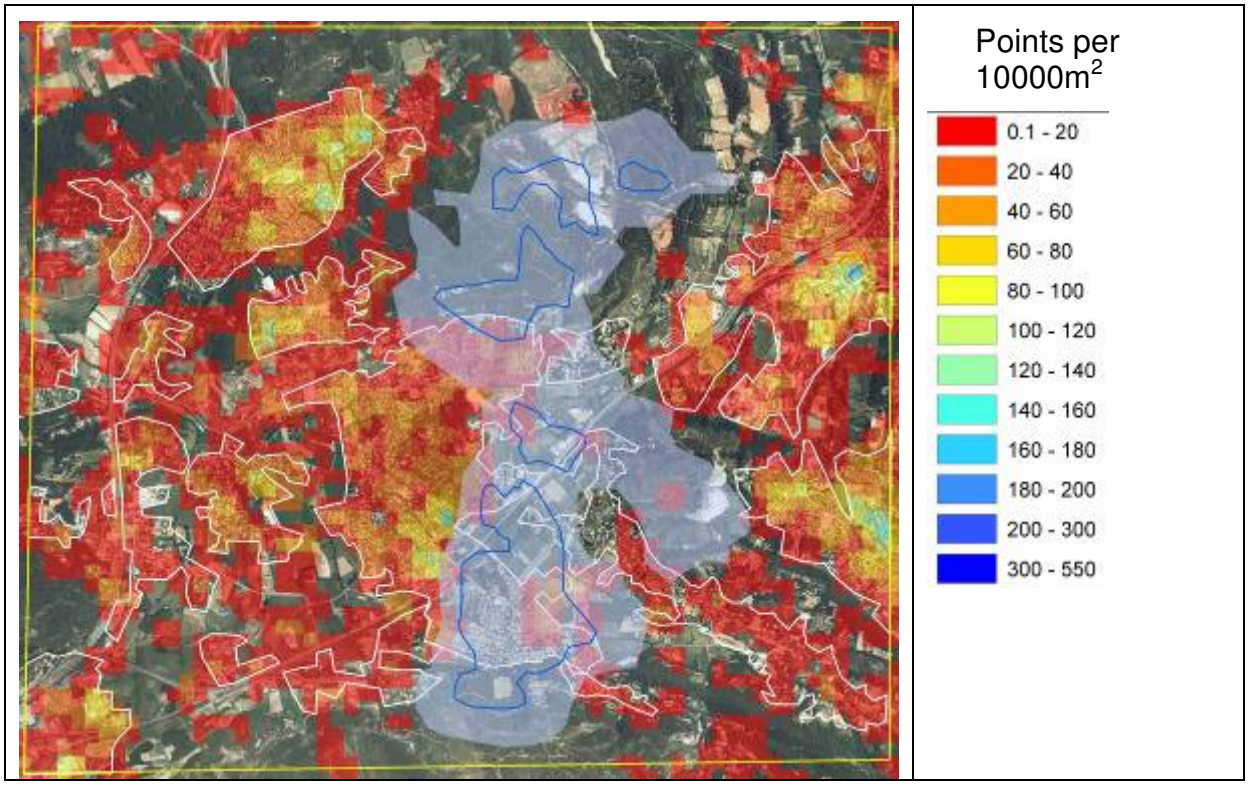


Figure 5.1.6 Team 3 densities. 10000m2 density; blue areas are high PS density within the team 3 dataset; red are lower density areas. Clear areas have no PS points. The yellow rectangle is the area defined for the overall density measurements. White polygons are urban areas, blue and red polygons are areas of subsidence, with red areas moving more rapidly that the blue areas.

Team 3 have a high density within the urban areas, but not many PS points in the more remote rural areas, points have been identified in the rural areas between urban areas; these would be areas where buildings do exist but would be wider spaced. They have picked up few PS points within the areas of rapid subsidence, even in the southernmost red area where many buildings are to be found.

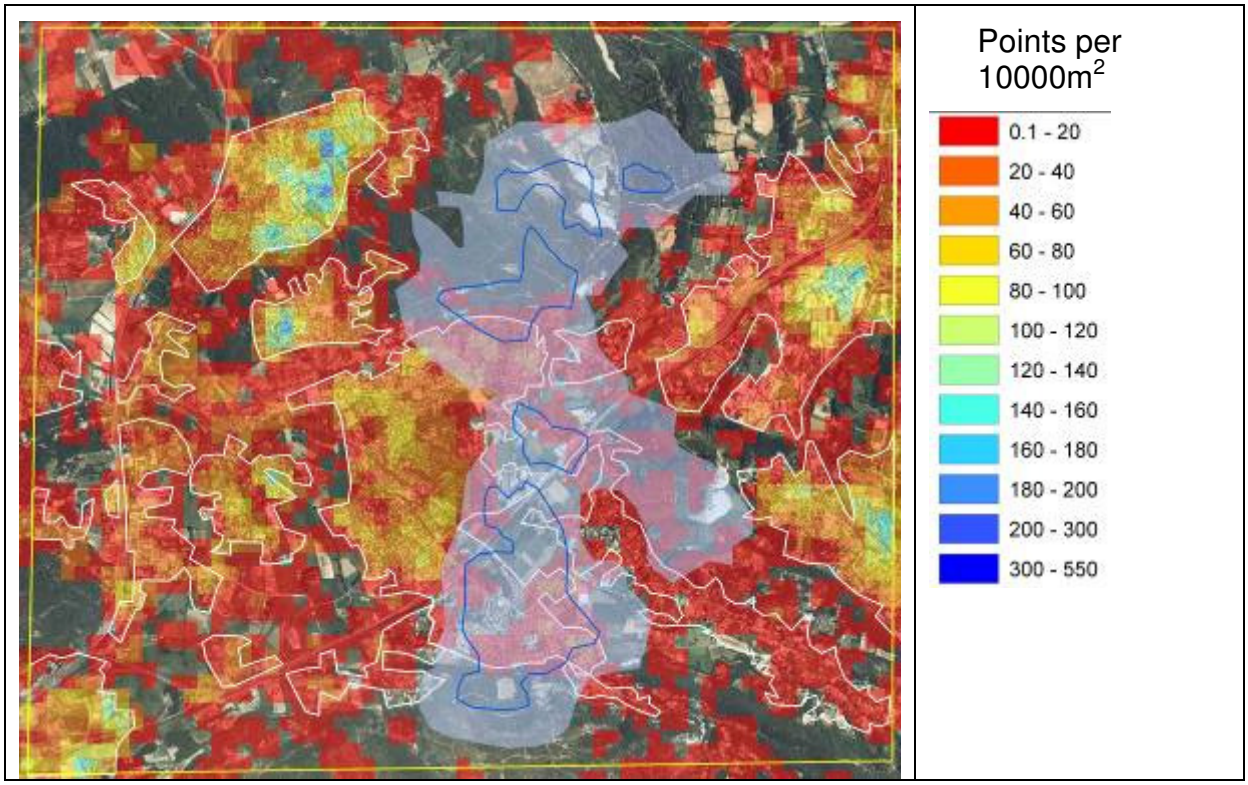


Figure 5.1.7 Team 4 densities. 10000m2 density; blue areas are high PS density within the Team 4 dataset; red are lower density areas. Clear areas have no PS points. The yellow rectangle is the area defined for the overall density measurements. White polygons are urban areas, blue and red polygons are areas of subsidence, with red areas moving more rapidly that the blue areas.

Team 4 found most of their points within the centre of the urban areas. Much like every other team, they also have a number of points within the rural areas that lie close to the urban areas. They have a limited number of points within the areas of rapid subsidence. Those that they did find, are closer to the edge of the polygon where subsidence is likely to be slower.

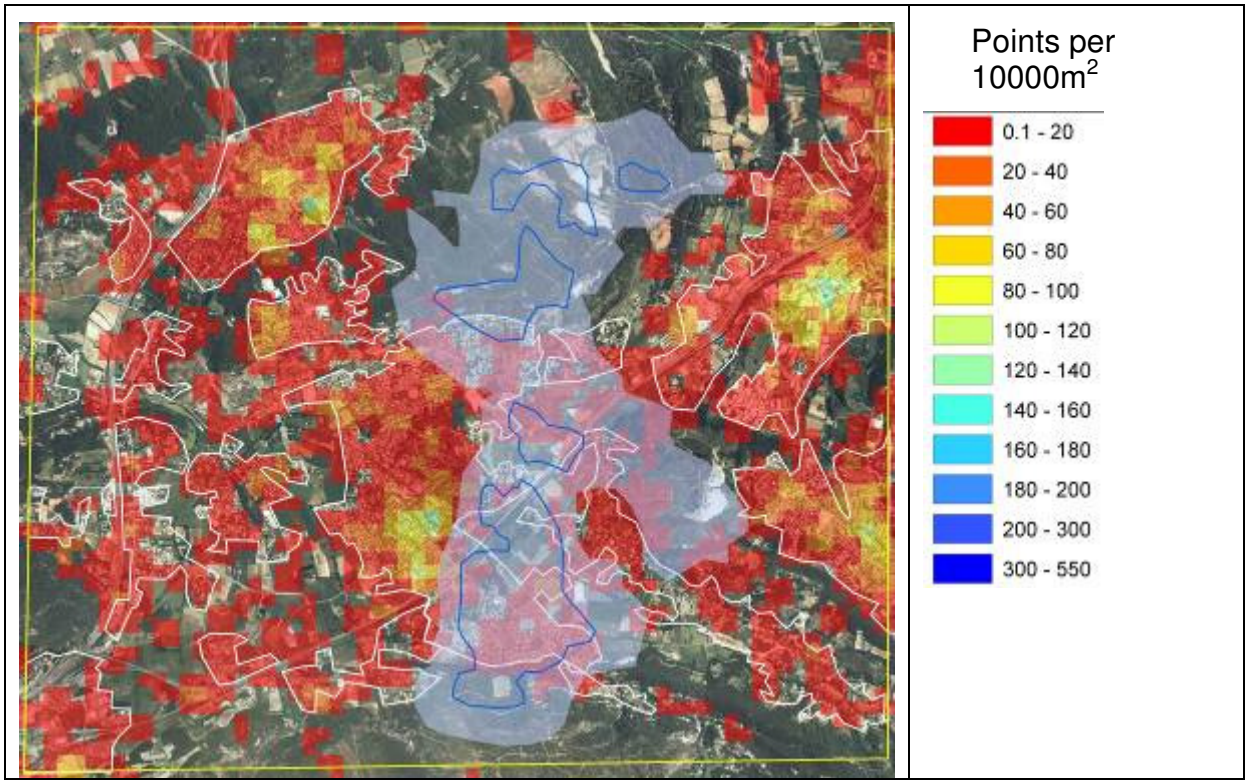


Figure 5.1.8 Team 5 densities. 10000m2 density; blue areas are high PS density within the Team 5 dataset; red are lower density areas. Clear areas have no PS points. The yellow rectangle is the area defined for the overall density measurements. White polygons are urban areas, blue and red polygons are areas of subsidence, with red areas moving more rapidly that the blue areas.

Team 5 has the highest densities within the urban areas. They do not seem to have picked up many points on the rural fringe of the urban areas, as other teams have. They have
some in the rapidly subsiding area to the south, but not many.

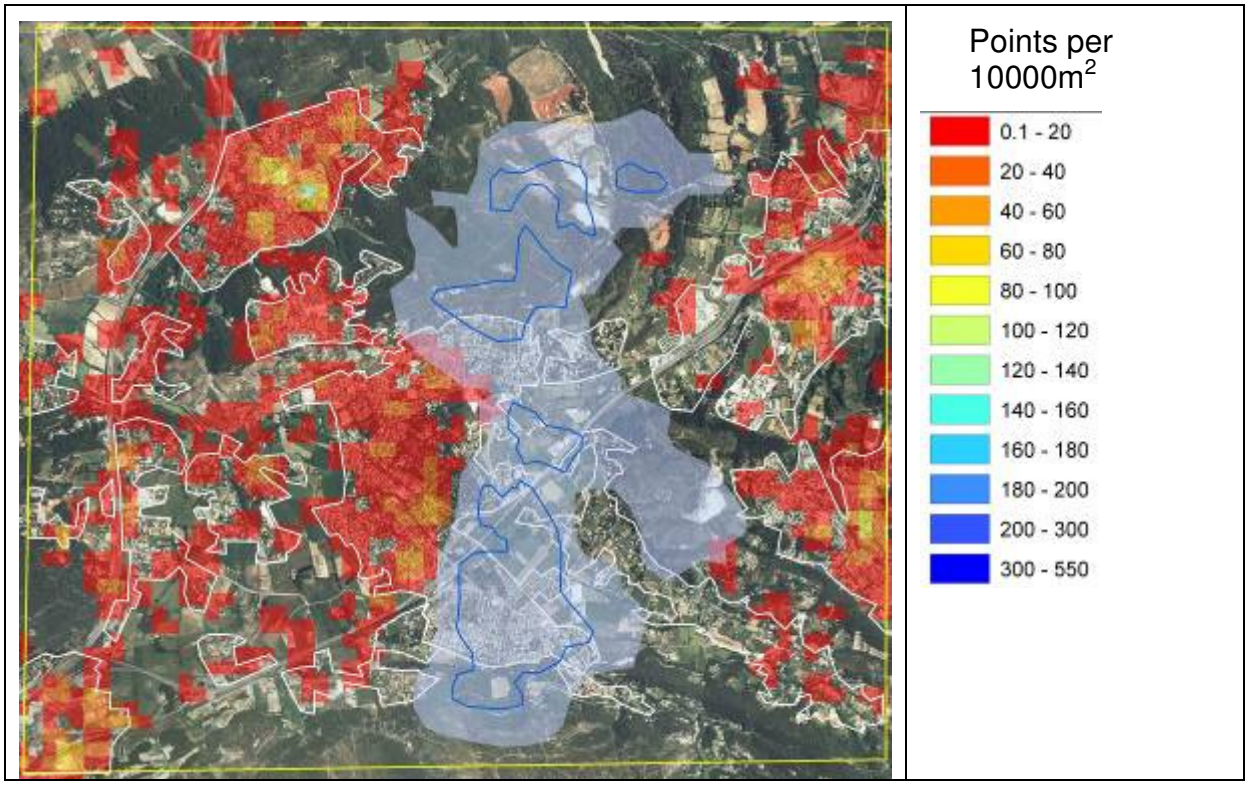


Figure 5.1.9 Team 6 densities. 10000m2 density; blue areas are high PS density within the Team 6 dataset; red are lower density areas. Clear areas have no PS points. The yellow rectangle is the area defined for the overall density measurements. White polygons are urban areas, blue and red polygons are areas of subsidence, with red areas moving more rapidly that the blue areas.

Team 6 found most of their points within the centre of the urban areas. However they have areas within the urban area polygons that are clear, suggesting no PS points at all. This may suggest that they are missing PS points that they should really be picking up. This fact is also illustrated by the fact that they have no points in the subsiding areas. Not only are they not detecting points in the red, rapidly subsiding areas, but also in the blue areas of less subsidence. They have no points in the rural areas.

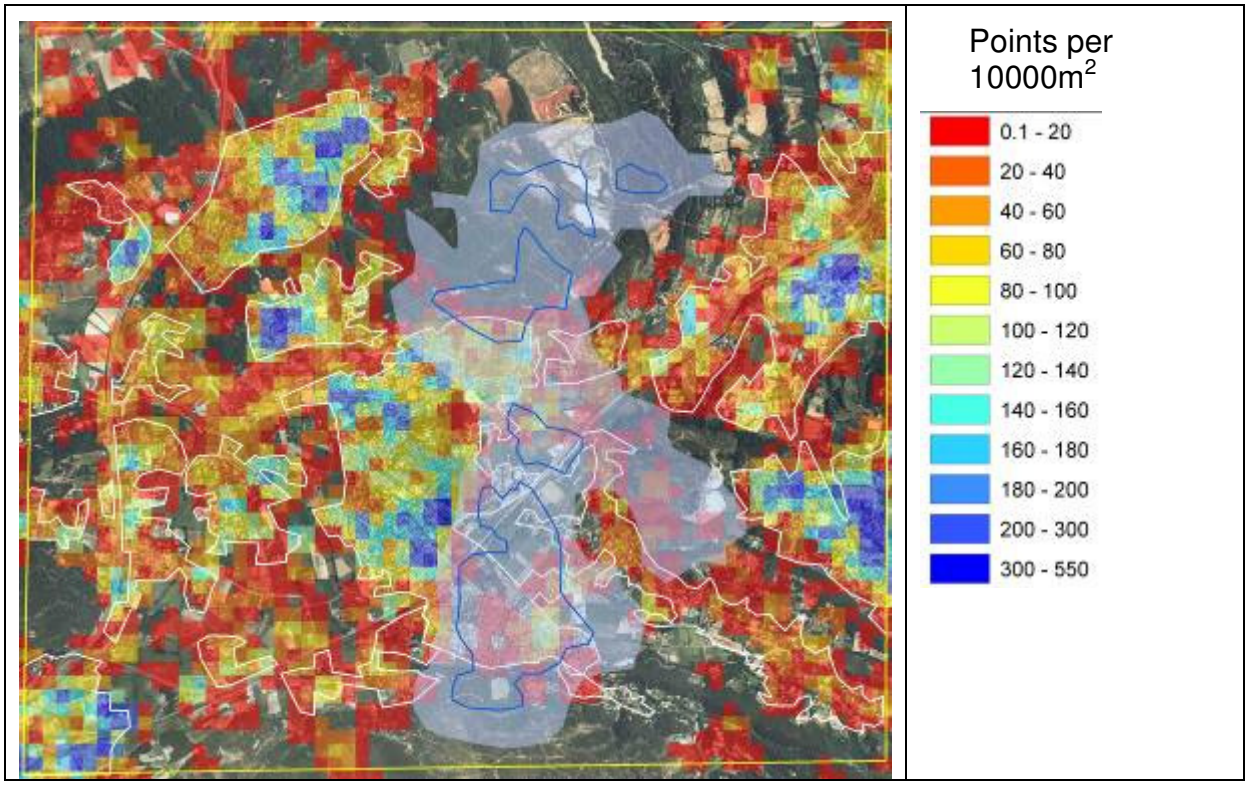


Figure 5.1.10 Team 7 densities. 10000m2 density; blue areas are high PS density within the Team 7 dataset; red are lower density areas. Clear areas have no PS points. The yellow rectangle is the area defined for the overall density measurements. White polygons are urban areas, blue and red polygons are areas of subsidence, with red areas moving more rapidly that the blue areas.

Team 7 found most of their points within the centre of the urban areas and some in the rural areas with urban areas in close proximity. They have few points within the rapidly subsiding areas to the south.

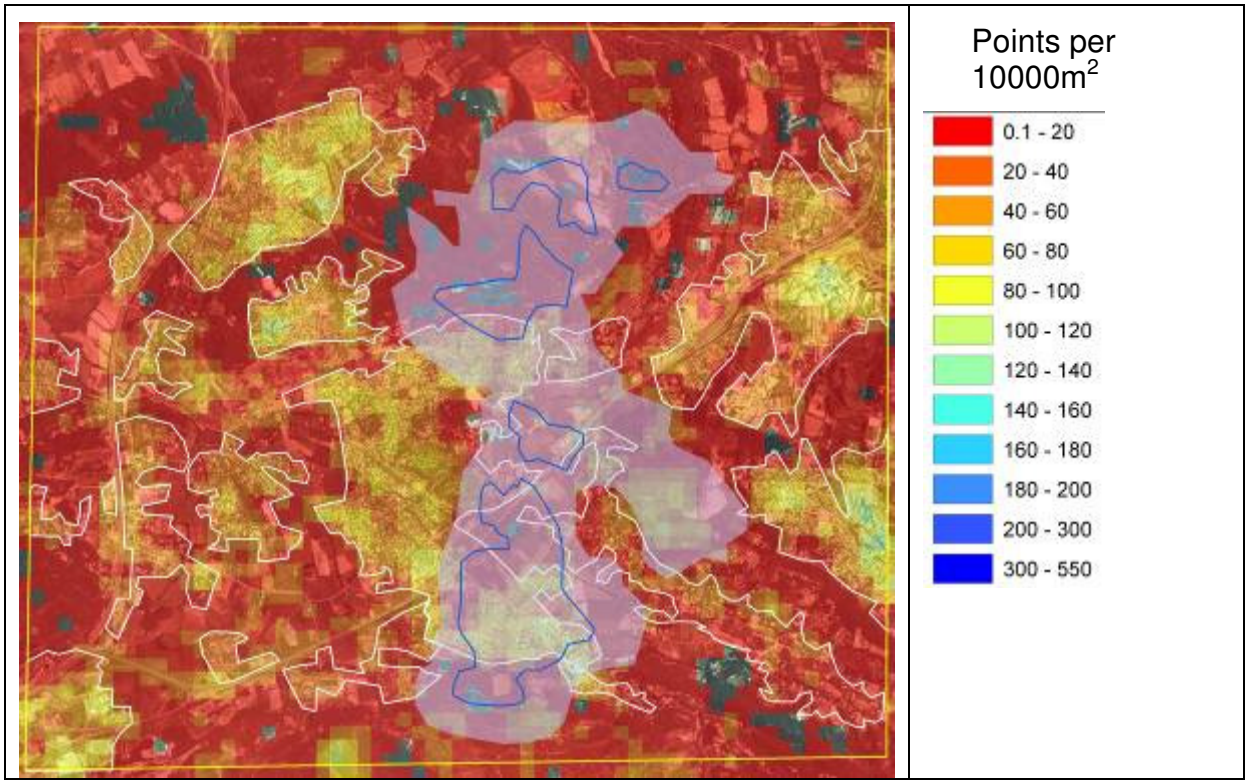


Figure 5.1.11 Team 8 densities. 10000m2 density; blue areas are high PS density within the Team 8 dataset; red are lower density areas. Clear areas have no PS points. The yellow rectangle is the area defined for the overall density measurements. White polygons are urban areas, blue and red polygons are areas of subsidence, with red areas moving more rapidly that the blue areas.

Team 8 found PS points almost everywhere. The highest density is within the urban areas. But they do have a few high-density cells within rural areas. However the measurement of the shifts has shown that these points are usually difficult to relate to ground features. Team 8 managed to detect a high density of points over the rapid subsidence area to the south, and even some points in the areas of rapid movement to the north, where the land use is rural.

It is also very interesting to examine the relative density distribution of the coherence values for all the PS points provided by each of the 8 teams. These are illustrated in Figure 5.1.12 below. The relative density is an estimation of the probability density function. It is similar to a frequency histogram, but smoothed, not binned. The range of the coherence values were divided into 200 bins at intervals of 0.05, and the mean coherence computed for each bin. Then a weighted density estimate was computed on the bin means, with the weighting based on the bin counts.

Teams used different definition of coherence. Therefore, a direct comparison between coherence distributions described hereafter would be meaningless. Although that, those distributions can be informative in respect to the PS selection process, which is based on coherence thresholding.

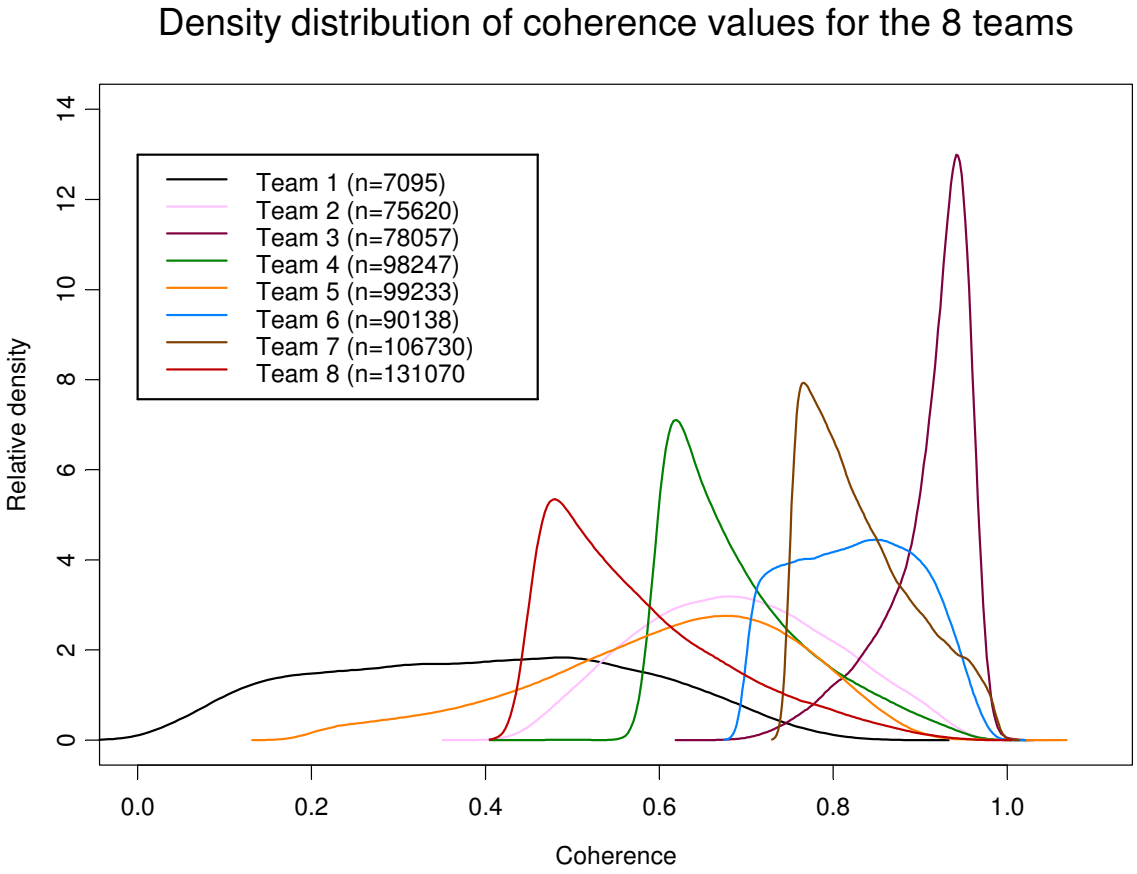


Figure 5.1.12 Plot of relative density distribution of the coherence values for the 8 teams.

These differences in the relative density distributions of coherence values might go a long way in providing some clues as to why the teams are giving different results in this study. Either they are using different definitions in calculating the coherence, or using different selection criteria right at the beginning of the processing chain on which PS points will be included in the analysis.

It can be seen that Team 1 has a very broad range of low to medium coherence values, with low relative density, and no well-defined mean. Teams 2 and 5 have similar shaped distributions with similar means, but Team 5 has a longer tail of lower coherence values. Team 6 seems to have an asymmetric bimodal distribution with medium and high coherence values. Teams 8, 4 and 7 have similar shaped relative density distributions of coherence, but with successively higher lower cut-off levels and increasing means, and are skewed to the right with higher coherence values. Team 3 has the highest coherence values, and narrowest peak of high relative densities, with a skewed tail to the left of lower values.

Conclusions on Spatial Distribution and Density

All teams did not process the same sized areas. However the size of area processed, does not seem to be the overriding factor on PS density detected. The teams that found the highest overall densities, for the defined measurement area, also found the most PS points for the urban areas.

As expected, the highest densities relate to urban areas, where permanent scattering objects exist. Surprisingly, some teams were able to find permanent scatterers in areas of forest or agriculture. Some teams (such as Teams 1 and 8) did succeed in finding PS points where

the permanent scattering features exist (such as urban areas) but ground motion was rapid. Other teams were not able to identify as many points in these areas.

In order to meaningfully compare the results of the teams in the rapidly subsiding areas it is necessary to compare the actual number of permanent scatterer points found within these areas. The following Table 5.1.2 shows the number of points each team found, within the area of most rapid ground motion, as identified from the DiffSAR results (also shown as the red polygons on Figures 5.1.4 – 5.1.11).

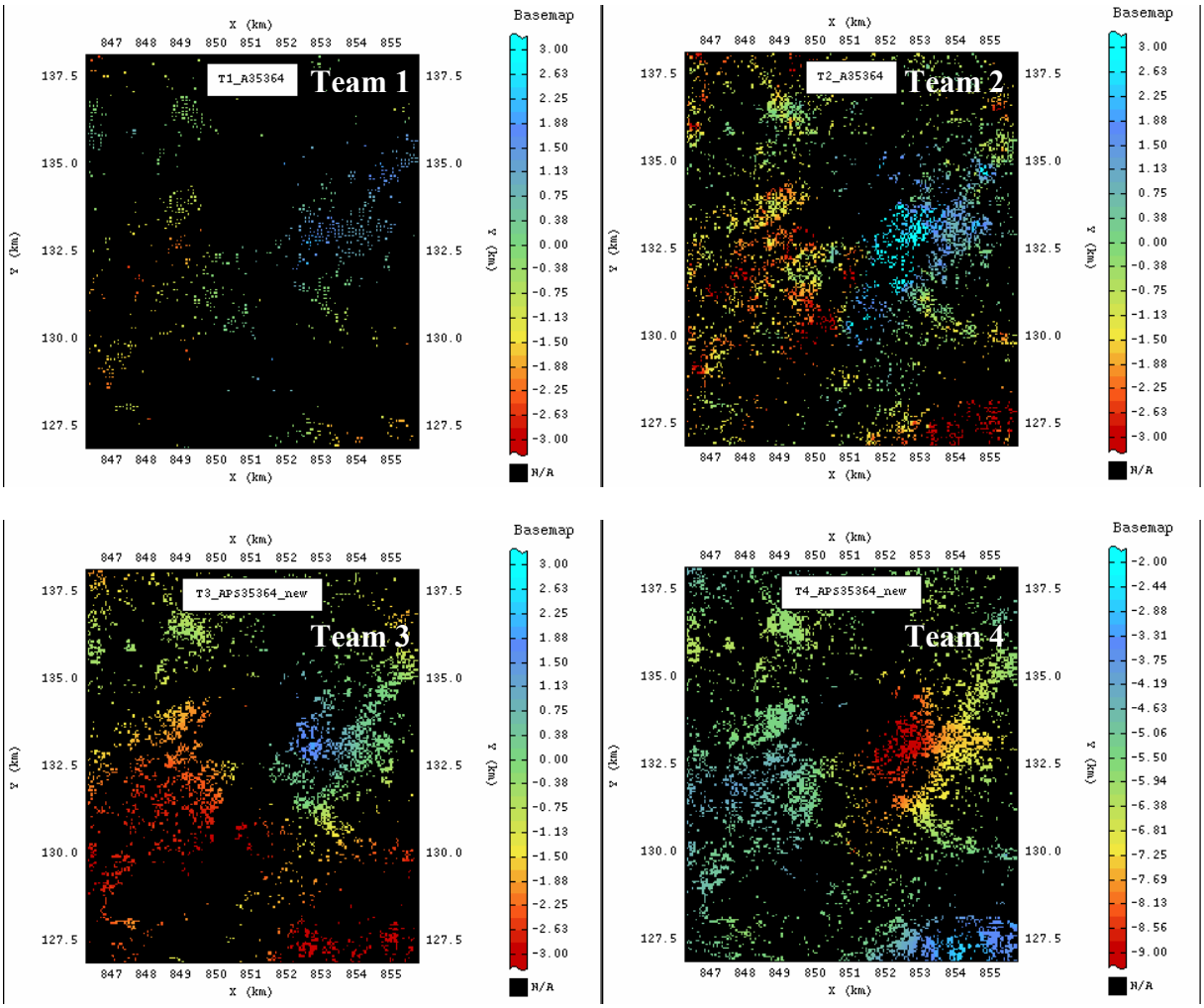
Team	No points in red subsidence area
1	32
2	105
3	69
4	90
5	92
6	0
7	138
8	1025

Table 5.1.1 Number of PS points which each team found in the area of maximum subsidence from the DiffSar results.

This table shows that Team 8 found the most points, by a large margin, within the area of subsidence. It also shows that Teams 2, 3, 4, 5 and 7 found a similar number of points. These teams also have similar looking density plots as shown above. Does this indicate that they are using a similar processing methodology? Team 1 and Team 6 would appear to be using different methodologies completely. Interestingly, Team 2 has a relatively high density in the subsidence areas but one of the lowest overall densities. Does their processing methodology allow them to detect a higher proportion of their overall PS points in areas of high motion?

5.2 APS intercomparison

We show below the APS plots from team 1-8 (except for team 5 who did not provide us with APS values) for date 26/10/96. APS plots should have *a-priori* the same trend of values considering that the phase delay due to the atmospheric path should be independent of the algorithm employed. Nevertheless, APS data represents not only the delay of the phase as the radar beams cross the atmosphere, but also contains un-modelled phase delays from different sources such as gradients due to un-modelled orbital fringes and phase components due to un-modelled non-linear deformation. As these residuals correspond to what was rejected to obtain the final deformation assessment a certain similarity between APS is expected; but one can notice from the plots below, each team has a different approach to APS modelling and presents different results.



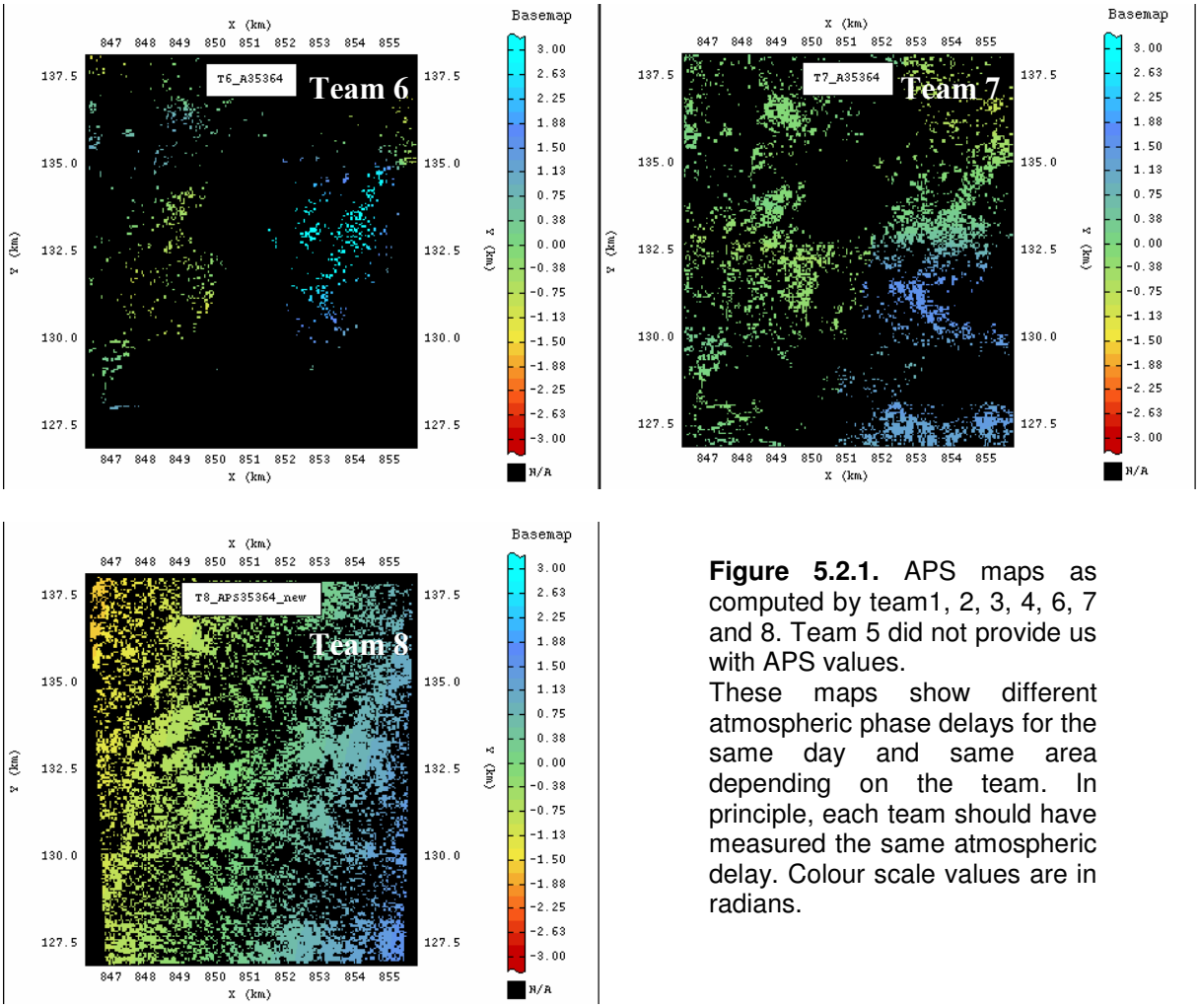


Figure 5.2.1. APS maps as computed by team1, 2, 3, 4, 6, 7 and 8. Team 5 did not provide us with APS values.
These maps show different atmospheric phase delays for the same day and same area depending on the team. In principle, each team should have measured the same atmospheric delay. Colour scale values are in radians.

5.3 Geocoding intercomparison

For each set of two teams A and B those grid cells in the radar image are selected having a radar signal detected by each team A and B. The following step is to calculate for each selected grid cell the distance between the respective team A and B Lambert coordinates. The average is taken as the 'average geocoding difference'. This procedure is followed for each set of two teams and results in the following table of the average of the geocoding differences in metres:

	1	2	3	4	5	6	7	8
1		55	46	69	38	69	45	70
2	*26*		15	20	16	41	17	37
3	*20*	*15*		16	6	44	7	39
4	*28*	*17*	*14*		29	42	19	36
5	*18*	*15*	*5*	*18*		48	10	43
6	*18*	*13*	*8*	*10*	*8*		43	23
7	*26*	*17*	*3*	*16*	*6*	*12*		41
8	*21*	*17*	*13*	*19*	*13*	*15*	*18*	

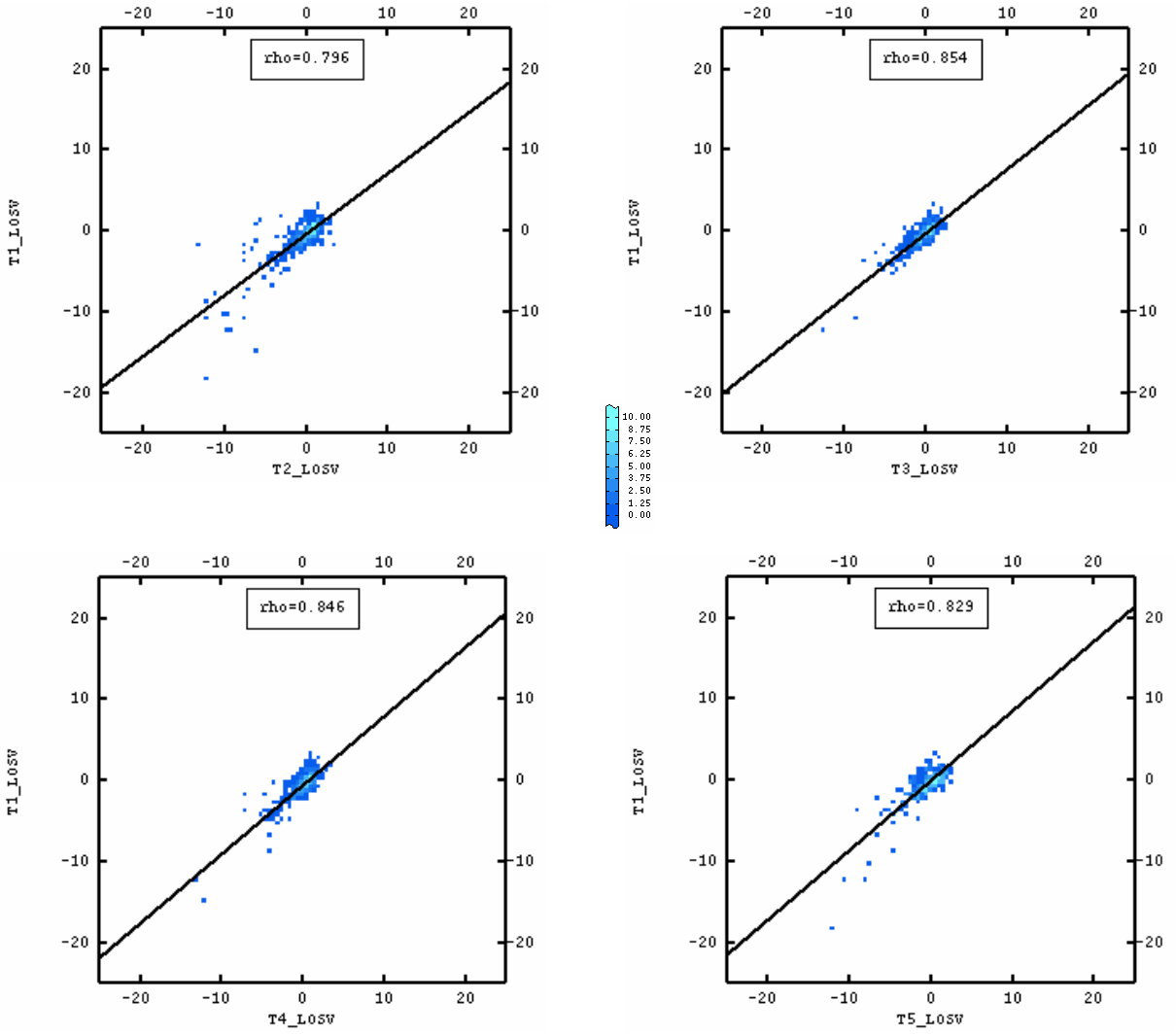
Table 5.3.1. Average geocoding difference (in m) and standard deviation (in m, notation in italics)

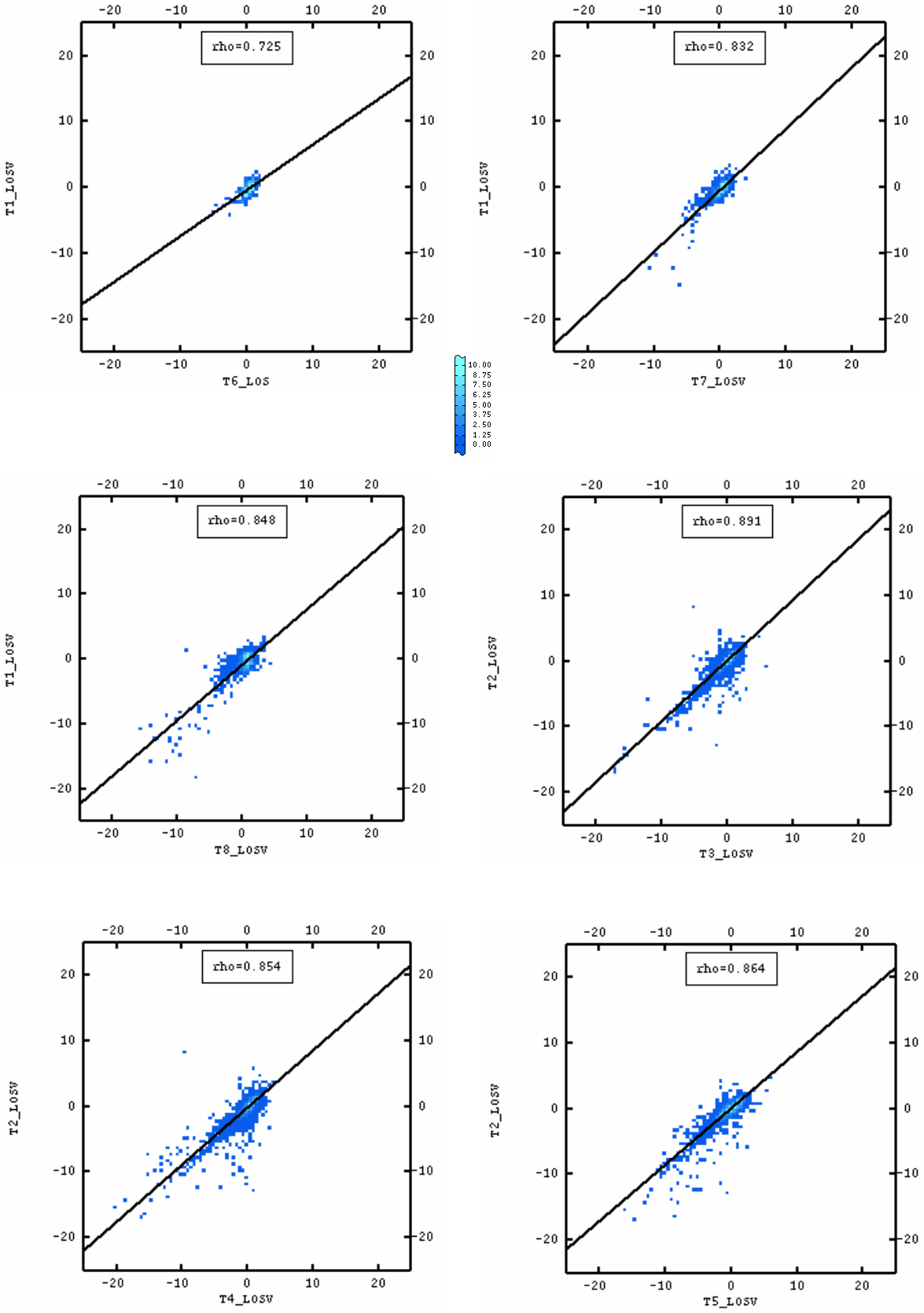
In principle the difference in geocoding should be zero. However, in practice there can be a difference of a couple of metres. Processing Chains 3, 5 and 7 have their radarcoordinates geocoded with less than 10 metres difference. Processing chains 2 and 4 have their radarcoordinates geocoded with 20 metre difference and differ on average by this amount with processing chains 3, 5 and 7. Processing chains 1, 6 and 8 differ by more than 20 metre with the other processing chains.

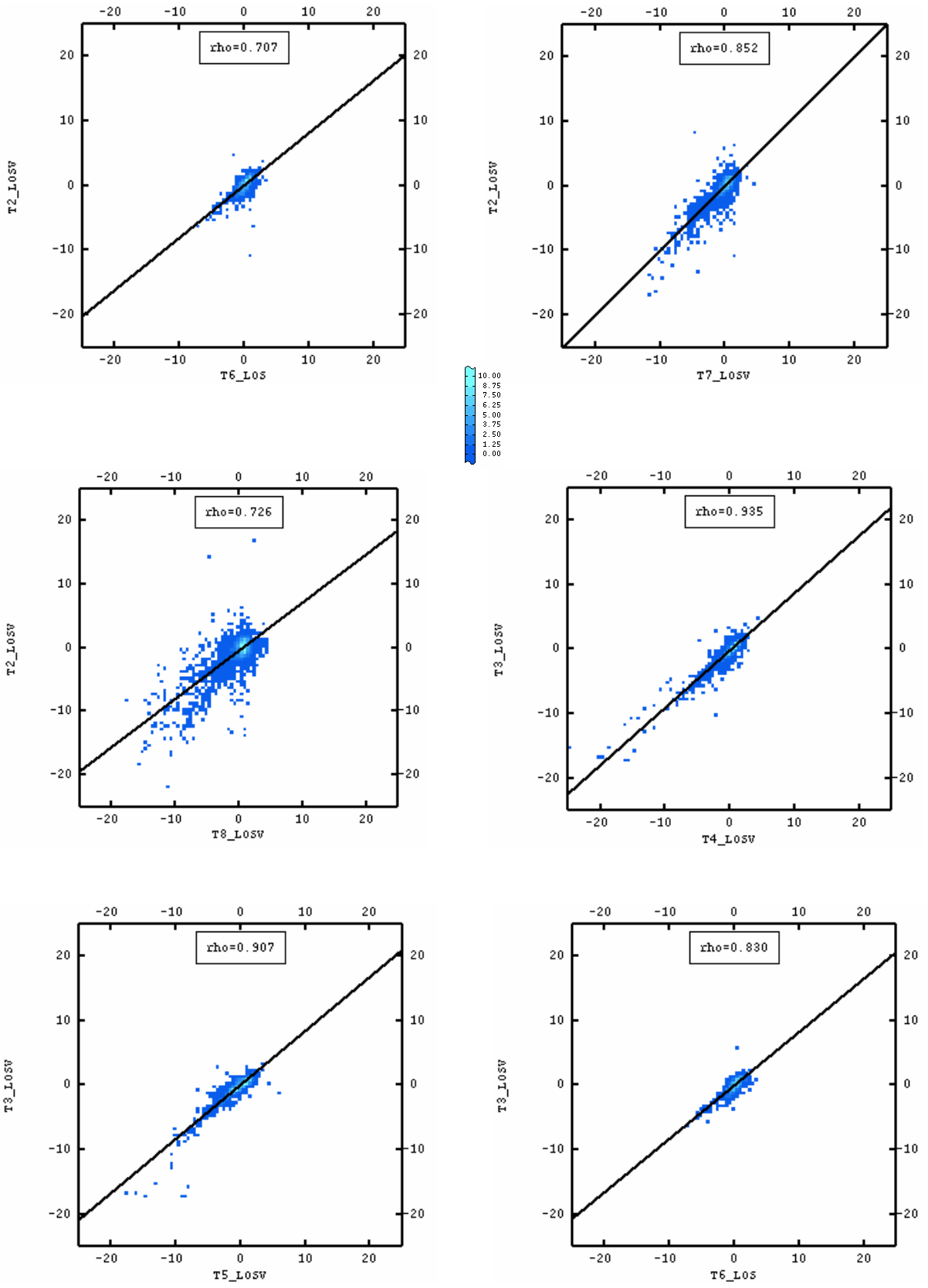
5.4 Intercomparison of velocity maps

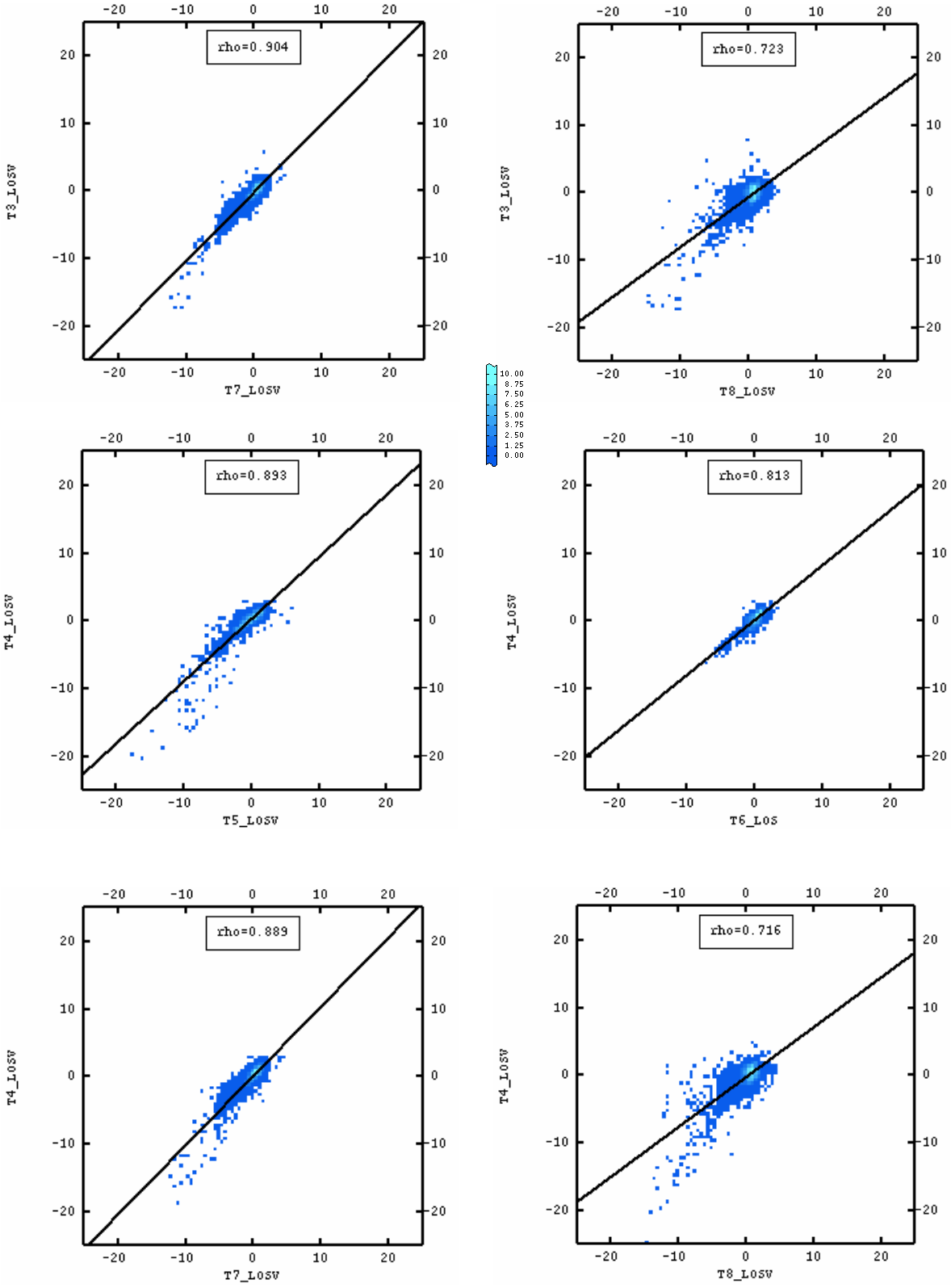
We present the intercomparison of LOS velocities. This intercomparison has been performed on velocities measured on a subset area, which all the teams have in common. Since team 8 processed the smallest subset of the SAR frames, the intercomparison has been performed on the area imposed by the Team 8 processed area (see 5.1. *Spatial distribution and density*).
We proceeded as follows; we built a grid of 50x50 metres on the selected intercomparison area. Then, we averaged the LOS velocities within each cell of the grid. Afterwards, we compared LOS velocities by pairs of teams on common cells only. The intercomparison is performed on cells within which the re is at least one PS. The scatterplot are shown below.

This task involved the comparison of velocity maps by pairs of teams. The hypothesis behind this comparison states that if each team had measured the same ground deformation rate over the study period, then velocity values would lie on a straight line, x=y, when velocity values of Ti are plotted against Tj. The graphic results are shown below in figure 5.4.1 while table 5.4.I goes over inter-comparison statistics.




rho=0.725
T1_LOSV
T6_LOS
rho=0.832
T1_LOSV
T7_LOSV
10.00
8.75
7.50
6.25
5.00
3.75
2.50
1.25
0.00
rho=0.848
T1_LOSV
T8_LOSV
rho=0.891
T2_LOSV
T3_LOSV
rho=0.854
T2_LOSV
T4_LOSV
rho=0.864
T2_LOSV
T5_LOSV


rho=0.707
T2_LOSV
T6_LOS
rho=0.852
T2_LOSV
T7_LOSV
10.00
8.75
7.50
6.25
5.00
3.75
2.50
1.25
0.00
rho=0.726
T2_LOSV
T8_LOSV
rho=0.935
T3_LOSV
T4_LOSV
rho=0.907
T3_LOSV
T5_LOSV
rho=0.830
T3_LOSV
T6_LOS


rho=0.904
T3_LOSV
T7_LOSV
rho=0.723
T3_LOSV
T8_LOSV
10.00
8.75
7.50
6.25
5.00
3.75
2.50
1.25
0.00
rho=0.893
T4_LOSV
T5_LOSV
rho=0.813
T4_LOSV
T6_LOS
rho=0.889
T4_LOSV
T7_LOSV
rho=0.716
T4_LOSV
T8_LOSV

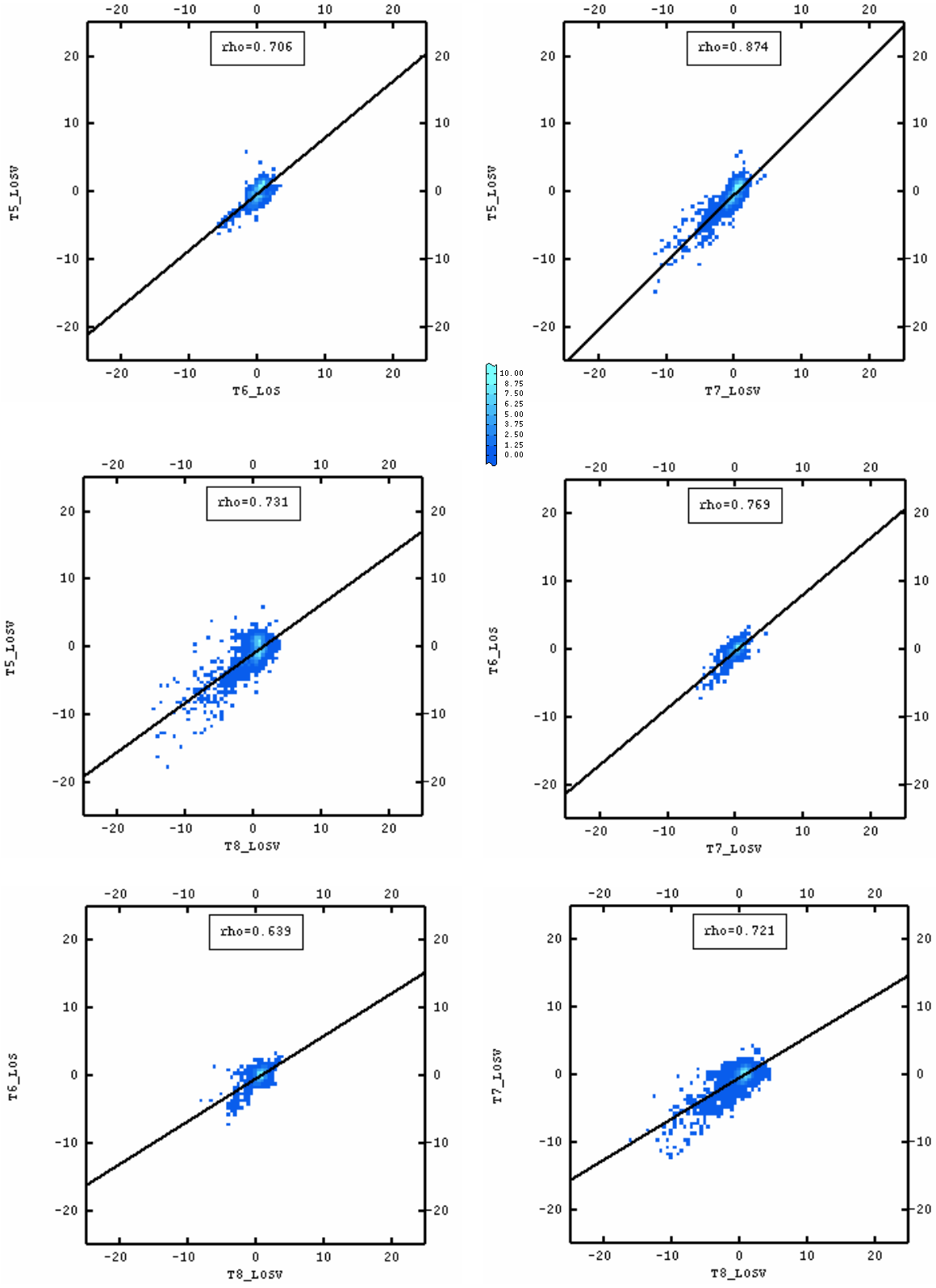


Figure 5.4.1. intercomparison of velocity maps. Scatterplots showing LOS velocity of Ti against Tj. The colour scale represents data density while the black line is the least square fit indicating degrees of data correlation. Axes units are mm/yr.
Data are concentrated around zero value and then spread out as velocity increases resulting in decreasing correlation between values as LOS velocity increases. All the teams agree with each other in measuring very small deformation rates (0 to 1.5 mm/y) then disagree when measuring higher deformation rates.

This exercise is meant to be a relative assessment as teams are compared by pairs one relative to another one. The LOS velocity intercomparison highlights discrepancies of the order of 0.05 to 0.8 mm/y in the velocity values retrieved by each team.

As shown in fig 5.4.1, data density is high and well localised around zero deformation rate. Data spreads out as deformation rate increases. Values dispersion (i.e. the radius of the cloud of points in each scatterplot) can be as small as 3 mm/y and as large as 9 mm/y depending on the team being compared. Velocity values are not well linearly correlated showing biases if put in relation with the slope of the x=y (perfect correlation) line. This is an important point in the intercomparison as teams are supposed to measure the same relative rate of deformation.

This intercomparison is not based on levelling data and therefore aims to asses the relative accord between independent observations. Consequently, disagreement in the LOS velocity values between pairs of teams suggests that LOS velocity retrieved by PSI technique rather depends on the measurement strategy adopted by each team.

STDV MEAN of Ti-Tj Correlation	T1 (889pts)	T2 (5100pts)	T3 (4169pts)	T4 (6552pts)	T5 (2511pts)	T6 (1269pts)	T7 (6360pts)	T8 (17081pts)
T1		1.57(0.79)	0.85(0.85)	0.88(0.85)	1.23(0.83)	0.84(0.72)	0.96(0.83)	1.44(0.85)
T2	-0.26		0.87(0.89)	1.06(0.85)	1.19(0.86)	0.99(0.71)	1.01(0.85)	1.86(0.73)
T3	-0.14	0.12		0.63(0.94)	0.87(0.91)	0.73(0.83)	0.71(0.90)	1.40(0.72)
T4	-0.46	-0.13	-0.24		1.01(0.89)	0.74(0.81)	0.76(0.89)	1.46(0.72)
T5	0.11	0.27	0.20	0.44		1.08(0.71)	0.99(0.87)	1.71(0.73)
T6	-0.34	0.06	-0.05	0.20	-0.21		0.81(0.77)	1.06(0.64)
T7	-0.32	0.02	-0.1	0.14	-0.32	-0.10		1.33(0.72)
T8	-0.82	-0.41	-0.56	-0.25	-0.84	-0.47	-0.40	

KEYS: STDV - MEAN of Ti-Tj – CORRELATION

Table 5.4.I. Mean and standard deviation of $Ti_{LOS} - Tj_{LOS}$ in red and blue respectively. The table also reports the number of points involved in the computation (between brackets) and the degree of correlation between pairs of teams. Values are in mm/y.
In the theoretical case the mean and standard deviation in this table should be zero. On the contrary, even if values of the mean range from 0.02 to 0.84 mm/y, the standard deviation is higher than the mean for each mean indicating that data are spread over a large interval. For instance, team 2 mean values are pretty low (e.g. 0.02) though the standard deviation values are as high as 1.01.

6. Discussion and Conclusions

6.1 Context remind

The PSIC4 exercise has been conceived as an inter-comparison and validation of PSI data computed by eight different teams. The aim of this exercise is to evaluate performances of eight different PSI methodologies by comparing the results with ground based observations and by inter-comparing the results from the eight processing chains. The results of this study provide an assessment of the absolute (validation) and relative (inter-comparison) accuracy of the PSI techniques. The PSIC4 exercise is a blind test, carried on a mining site characterised by different magnitudes and evolutions of deformation and variable land cover. It is worth emphasising that the PSIC4 represents a unique experiment for the number of PSI teams involved and for the quality of the available ground truths, which involves more than 1000 levelling benchmarks measured once or twice per year. The project involved an important validation effort, which included the preparation of a detailed validation plan, different pre-processing steps, and several validation and inter-comparison activities. Furthermore, a special effort was devoted to a comprehensive cross-check of the key activities and tasks of the project. This cross-check has shown that the outcomes of the validation and inter-comparison results are consistent and reliable: they represent a technically sound base to be used to assess the PSI performances over the considered test site.

The PSIC4 project therefore represents a valuable validation exercise. However, it is important to underline that the deformations of the PSIC4 test site, especially those of the mining area of Gardanne, where the majority of the ground truth are located, represent a difficult case for the PSI techniques. In fact, the deformations range from few centimetres up to some decimetres, and most of the deformation occurs in few months, a rather short period for the (at most) monthly SAR acquisitions. Note that the above deformations are “strong deformations” from the viewpoint of C-band PSI with the acquisition capabilities of ERS and Envisat, which represents the context of this project, and cannot in principle extended to other types of configurations, e.g. L-band or more frequent SAR acquisitions. Therefore, in PSIC4 the performance of PSI techniques was tested on the very edge of its capability, as the PS interferometry processing is performed in the less favourable conditions and is evaluated upon the strongest criteria.

When going through the results of the PSIC4 exercise, one should ever keep in mind the following issues in order to interpret well the outcome of the study:

1. The PSIC4 exercise is conceived as a blind test in order to assess the applicability of PS-InSAR technique. The eight teams did not have knowledge of the type of deformation occurring on the test site, i.e. information such as the linearity/non-linearity of the deformation, the driving mechanism, the location of the deformation, when it started and when it ended, and expected deformation magnitude. Moreover, the eight teams were not aware of which spatial subset of their data would be validated and compared after the PS processing (*i.e.* whether the validation would be performed on a relatively stable area or on a deforming area). More in general, the teams received no information on what exactly was the deformation signal of interest,

i.e. it was not clearly specified the goal of the PSI analysis. By contrast, the validation was focused on a specific deformation phenomenon, i.e. the deformation associated with the mining area of Gardanne. This point is important because the PSI data processing has different parameters that can be tuned and tailored to a specific application goal. For instance, the processing can be modified to take into account a priori knowledge of fast displacements. This clear advantage of a tuneable PSI processing was not used in the PSIC4 project. As a consequence, each team was left free to deliver PS data points where they judged them the most reliable, getting for instance different densities of the measured PS.

2. Each team has processed a different spatial subset of the SAR frames. Therefore, the spatial coverage of the processed data changes from team to team. This issue affects the intercomparison process (in principle it does not affect the validation) as the intercomparison had to be done compulsorily on the common processed area (*i.e.* the smallest area that teams have in common).

3. The PS measurements were evaluated upon the strongest criteria. For instance, for the first time the validation of PSI against levelling data is performed quantitatively to the millimetre level. Some of these criteria can be non relevant in specific applications. This for instance can concern the geocoding accuracy, which in large scale deformation studies may not represent a strict requirement but that could be needed in many applications

These introductory remarks illustrate that whatever the PSIC4 results are, the PS technique is not invalidated; instead, the outcomes of the PSIC4 project should represent an input to improve PS interferometry performances over the critical application cases.

6.2 Summary of the presented validation and intercomparison results

6.2.1 Pre-processing

Before starting with the validation/inter-comparison activities, a series of pre-processing steps have been carried out

Correction of geocoding.
We observed shifts ranging from 5 m to 85 m.. A correction was applied when possible to allow the validation and intercomparison procedures.

Spatial resampling of the PS and temporal resampling of levelling. Since PS-scatterers and levelling points do not tend to co-register, a spatial resampling has been done of each of the PS-datasets, thus obtaining a best estimate of PS-derived displacement time series at the location of the levelling points. The levelling was resampled to provide deformation values for the benchmark at the SAR acquisition date (SAR data is temporally denser than the levelling) and an interpolation error in order to have a confidence indicator on this value. The spatial resampling was further motivated by the fact that in the direct neighbourhood of a levelling location many

PS-scatterers can be found. To pick only one of these as representative for a comparison with the levelling time series would introduce a bias. Therefore, a spatial resampling has been performed using a geostatistical algorithm in order to get an unbiased estimate. A further test has been applied on a selection of individual PS-time series from which it was concluded that a 50 m sampling space around the levelling points did not introduce significant bias in the resampling.

Stable area correction. As the teams used different stable areas as reference, the PS time series have been corrected from velocity biases by selecting a stable area and subtracting the average velocity of this area.

6.2.2 Intercomparison

Spatial distribution. The highest densities relate to urban areas, where scattering objects exist. Surprisingly, some teams were able to find points in areas of forest or agriculture (Team 8 in particular). Some teams (such as Teams 1 and 8) did succeed in finding PS in urban areas and within a rapid ground motion zone. Other teams were not able to identify as many points in these areas. The fact that most of the teams rejected points in the area of maximal deformation is an important issue for the intercomparison (because otherwise the comparison is performed on the areas of moderate deformation) and in terms of PSI technique application: which are - in a general case - the characteristics of the deformation field preventing PSI technique to be applied? See the discussion of the Open questions in Appendix A. The PSIC4 exercise showed that for the case under consideration the main area of subsidence could not or only partly be assessed and identified by 7 out of 8 teams. The main reason for this has been the low density of permanent scatterers in the area of interest. Thus, masking the actual subsidence bowl in a way. Nonetheless, improvements are possible taking into account that one team did found a high distribution of Persistent Scatterers within the subsidence area.

Overlying with ancillary data The comparison with qualitative information is meaningful for the understanding of the PSIC4 issues since it might shed light on how the PSI-measurements relate to natural phenomena. All the teams that have PS points within the areas of a particular mined panels show subsidence trend in the time series. However, the PS subsidence on time series nearly always start before the mining activity took place on that particular panel. That shows a possible methodological bias. This could be related to the choice of the deformation models so that the partial step-like deformation profile of the active mining subsidence seems to be adjusted to a linear deformation model starting before deformation has occurred. We have to notice that Team 8 results are clearly more consistent with the expected temporal profiles of deformation.

APS-intercomparison. The purpose of this step was to compare the so-called Atmospheric Phase Screen (APS) of the different teams. Those APS in fact correspond to the obtained phases once the PS heights and deformation has been modelled and withdrawn. In some cases there is practically no correlation between APS from different teams. Moreover, the standard deviation of the differences is in most cases larger than 1 radian and up to 2.1 radians. Once converted in displacement errors at the acquisition date that would correspond to discrepancies between 4.5 mm and 9 mm.

Velocity intercomparison. Velocity is the basic deformation parameter derived from PSI techniques as it is obtained by assessing a linear regression on the phase history. A very high precision was expected. The inter-comparison results show discrepancies, in terms of standard deviation, between 0.6 to 1.9 mm/yr.

Geocoding comparison. Significant differences between teams with magnitudes of 'average geocoding difference' in between 6 and 70 metres.

6.2.3 Validation

Time Series. A comparison of the spatio-temporal profiles of the levelling data and PSI data along two levelling transects shows that for this case, the stretch along the line experiencing most subsidence was not well sampled by 7 out of 8 teams and an underestimation of subsidence velocity was shown by all teams. This finding can be explained by two facts. Firstly, part of the subsidence area is located in non-urbanized area. This precludes the finding of suitable scatterers and thus makes it impossible by the mere nature of the methodology to properly characterize the subsidence area. Secondly, it was found by comparing time series of resampled PS-scatterers and levelling that there is *a strong relation between RMSE and the magnitude of displacement inbetween SAR-images*. This fact strongly suggests that the underestimation of the subsidence (in space and time) has been the result of phase unwrapping ambiguities. A comparison of the PSI time series and levelling time series shows that if the displacement is larger than about 2 cm inbetween two consecutive SAR-images, PS-InSAR starts to seriously deviate from the levelling time series. Since for the Gardanne site, a large number of the levelling time series show large displacements for two consecutive SAR-acquisition dates (35% in excess of 2.8 cm and 70% in excess of 1.4 cm) validation results are therefore negatively biased. This also explains the low number of cases for which PS-InSAR time series and levelling time series could be tested to belong to the same population. If one would only compare those time series having a maximum displacement of 1,4 cm or less for two consecutive SAR-acquisition dates, than the current study *shows average RMSE between levelling time series and PS-InSAR time series of 7 mm to 25 mm*. One has to compare this with validation studies performed on artificial scatterers which show standard errors of the time series of 2 mm. From this it can be concluded that for those locations for which phase unwrapping ambiguities do not exist, at least some of the processing chains obtain results in line with previous studies which mainly took place under controlled circumstances. In all cases reviewed, the team's results did underestimate the subsidence rate in areas showing moderate to fast subsidence. The main advanced reason for this is the character of the subsidence process in the study area. As a result of mining activities in this area, the subsidence takes place in a relatively short time-span. The strong correlation between RMSE and magnitude of displacement for two consecutive SAR-acquisition dates, suggests that the results have certainly been affected by phase unwrapping ambiguities, leading to a systematic underestimation of the subsidence rates.

Velocities. The comparison of the PSInSAR velocity with the ones derived from levelling shows an average absolute difference with standard deviations between 5 and 7 mm/yr. However those standard deviations are strongly dependent on the

absolute value of the actual displacement of the measured point. It can reach more than 15 mm/yr on the main deformed area but generally less than 2mm/yr on stable levelling points.

Check of geocoding quality.

The task was a finer check (after shift correction) of the geocoding. The observed errors range from 1 m to 20 m. In most cases this final accuracy allows to locate the PS at a building base. We have to notice that 2 teams were not corrected, in particular the one with the bigger error value.

The following table summarizes the main results presented in the document.

Validation task	***Extracted indicator***	***Approximate observed values***	***Expected values***	***Section***	***comments***
Geocoding shifts	Shift values	5→85m	Ideally <10m (to be able to separate building)	2.3	T1,T8 : shifts were not reliably identifiable T6: shift > 70m limits the applicability of the result However, one has to take into account that the team 6 did not correct the geocoding with the ortoimage. Then, this observed value (85m) is not significant T3: shift in the expected accuracy. T2/4/5/7: shifts nearly acceptable
Time series validation	Visual comparison of spatio-temporal evolution of deformation along levelling lines	Qualitative observation	Similar behaviour compared with levelling but also between the different teams	3.2	Most teams identified a deformation, but failed in identifying the complete subsidence bowl and magnitude along the levelling lines in question due to low density of scatterers and phase unwrapping ambiguities.
Time series validation	RMSE respect to levelling	On average 7 mm – 25 mm for levelling series having a displacement inbetween SAR-acquisition dates smaller than 1,4 cm. For other time series RMSE can increase to > 200 mm.	no a-priori information	3.2	The discrepancies are higher than reported for current studies on stable corner reflectors. Exceptional high RMSE is found for fast subsiding locations most probable due to phase unwrapping ambiguities.
Time series validation	Student t-test of PSI-time series in relation to levelling time series to test the goodness of fit	Maximum about 10% of all time series with a displacement less than 2,8 cm inbetween SAR acquisition dates are tested to be from the same population.	no a-priori expectation	3.2	The discrepancies increase with the deformation values most probable due to phase unwrapping ambiguities.
Velocity validation	Mean of the absolute value of the global differences (to test the accuracy of the techniques)	2.5mm/yr → 4.5mm/yr	-	3.3	The observed values seems consistent with deformation > 5mm/yr
Velocity validation	Standard deviation of the global differences	5mm/yr → 7mm/yr	-	3.3	The observed values seems consistent with deformation of the order of the cm/yr

Velocity validation	Standard deviation differences in relation with the actual deformation	More than 15mm/yr for area with deformation larger than 15 mm/yr		3.3	That confirms the loss of accuracy in the faster deformation areas previously observed
Overlying with ancillary data	Observations in relation with the knowledge about the deformation (to identify inconsistencies in the temporal behaviour of the assessed deformation)	qualitative		3.4	The behaviour of the time series seems to not fully correspond to the expected in relation with the mining works: "linearization" step profiled deformation seems to affect most teams
APS inter-comparison	Correlation (to test the similarity of the residual)	0.0→ 0.92	nearly 1	4.2	Team 2,4,3,6 seem to have a certain similarity in the residual . The other teams have poor correlation with them.
APS inter-comparison	Standard deviation of the difference of APS	0.66→2.14 rad		4.2	Dispersion between teams ~ 1/5 fringe
Inter-comparison of velocity maps	Standard deviation of the difference of velocity (to test the similarity between the velocity products)	0.6→ 1.9mm/yr	Of the same order from the announced intrinsic precisions of the technique (0.1→0.3 mm/yr)	4.4	The discrepancy on this test (which takes into account only the PSInSAR data) are higher than we expected
Inter-comparison of velocity maps	Mean of the difference of velocity	0→ 0.8mm/yr (remark: a stable area correction has been applied)		4.4	Most teams' pairs have small mean of difference respect to their standard deviation

6.3 Final discussion

This section includes some of the key outcomes of the PSIC4 project and other complementary activities, whose results are collected in this final report. These activities are briefly described below.

The main outcomes of the validation and inter-comparison activities over the Gardanne test site have been described in the previous sections. Starting from these results, the PSIC4 validation team prepared a list of "Open questions" to be discussed during the PSIC4 Final Presentation and the PSI validation workshop, held in ESRIN the 18-19th September 2006. These questions were organized in two main categories: those directly related to the PSIC4 results, and more general questions related to the PSI techniques.

The responses to the "Open questions", collected during the PSIC4 Final Presentation and the PSI validation workshop, with additional contributions collected after these two events are included in the Appendix A of this Final Report. This document represents a valuable outcome of the PSIC4 project, because it contains

the technical explanation of most of the results achieved in Gardanne. Some of the key comments of this document are summarized in the following points, along with other relevant considerations.

1) One of the most important conclusions of the project concerns the characteristics of the mining test site of PSIC4, which include abrupt nonlinear movements with magnitudes that range from few centimetres up to some decimetres. These are severe characteristics from the viewpoint of C-band PSI with the temporal acquisition capabilities of ERS and Envisat. Why the deformation characteristics are so important? Because in principle PSI can measure surface displacements with millimetric precision, but this can be achieved under the following conditions:

- The right model to describe the deformation is adopted. This is difficult to accomplish with abrupt nonlinear movements.
- The aliasing due to low PS density and/or low temporal sampling with respect to deformation, which may cause phase unwrapping errors, is controlled. This is difficult or impossible with strong deformation magnitudes.
- The assumptions to separate the atmospheric contribution from deformation are correct. This typically fails in presence of nonlinear motion.

Most of the results of PSIC4 can be understood in the context of the above conditions: none of them is fully accomplished in the mining area of Gardanne.

2) It is worth to underline that the above consideration on “strong deformations” holds for C-band PSI with the current temporal acquisition capabilities of ERS and Envisat. They cannot in principle extended to other types of SAR missions based on different band and more frequent SAR acquisition capabilities.

3) The PSIC4 project was conceived in a specific framework, where the teams worked under “blind conditions”, with no a priori information on deformation type, driving mechanism, deformation magnitude, etc. Furthermore, they had no information on what exactly was the deformation signal of interest, i.e. the goal of the PSI analysis. By contrast, the validation was focused on a specific deformation phenomenon, i.e. the deformation associated with the mining area of Gardanne. This point is important because plaid a key role in the PSI processing. In fact, instead of tailoring the processing to a specific objective of the analysis, the teams used a standard approach and a processing which is feasible with reasonable resources. It is worth emphasising that the area covered by most of the teams is considerably larger than the 100 km^2 area used for the validation. Any of the PSI teams has performed an advanced or refined PSI analysis, because neither the area of interest nor the goal of the refinement were defined). This again explains most of the PSIC4 results, e.g. the lack of PS in the mining area “of interest”.

4) It is worth analysing a specific consequence of the above point, which explains the different densities achieved by the teams. The PS densities are different because there is no “definition” of what exactly is a PS. The teams used their standard PSI approach (instead of an advanced or tailored one) and delivered the PS only where both velocity and time series could be extracted with reasonable reliability. Unfortunately the validation area represents a difficult area, where phase unwrapping errors represent the main problem. Due to the high probability of this type of errors many teams did not deliver unreliable information. Note that this did not occur outside the mining area, i.e. in the great majority of the covered areas, see point six.

5) Considering the above points and the results achieved in the Gardanne mining area we can say that PSIC4 has clearly demonstrated that the PSI limitations are real, i.e. that PSI is not applicable everywhere. Though this was already clear to many PSI specialists, now this evidence has been widely documented.

6) The PSIC4 results only concerned the “difficult” mining area of Gardanne, while most of the PSI teams covered considerable larger areas, which include gentle deformation and stable areas. For this reason, within the ESA-founded project Terrafirma was decided to run a additional inter-comparison study over the same dataset of PSIC4. This work, named “Provence Inter-comparison” was restricted to four of the eight teams of PSIC4. The final report of “Provence Inter-comparison” is reported in the Appendix B of this final report. This report offers an interesting complementary source of results with respect to PSIC4. When comparing the results of the two projects, one may notice an increase of the performances in the case of the Provence inter-comparison for both the deformation velocity maps and the time series. It is worth to underline that the results of the two projects (Provence inter-comparison and PSIC4) are not contradictory. In some cases they simply show different complementary aspects. For instance, the Provence inter-comparison is largely based on data outside the Gardanne mining area, which were simply not analysed in PSIC4. In other cases the results are rather similar, see the example in the Conclusions of the Appendix B.

7) To conclude, the limitations of PSI over deformation areas with similar characteristics to Gardanne open a series of important issues for the future. The first one is the importance of a feasibility study before running a PSI analysis. This may help in avoiding false expectations and disappointing results. Note that a feasibility study is now implemented within the Terrafirma project. A second issue concerns the appropriate ways to inform the PSI users of the limitations of the technique, especially in those cases where PSI is employed under the non-ideal conditions. Then it is interesting to investigate the possibility of using alternative techniques to PSI, like DInSAR which could provide useful information in difficult applications like the mining areas.

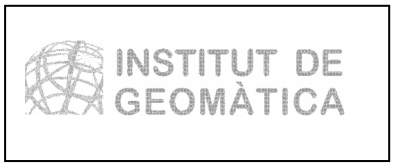
PSIC4

APPENDIX A
of the Final Report of PSIC4:

Open Questions

Michele Crosetto, Institute of Geomatics
Marcus Engdahl, ESA

Final Version
Castelldefels, 15 March 2007



PSI validation workshop, ESRIN 18-19 September 2006: Open questions

M. Crosetto, Institute of Geomatics
M. Engdahl, ESA

Introduction

The PSIC4 Final Presentation and the PSI validation workshop, held in ESRIN the 18th and 19th September 2006, were devoted to the presentation of the final results of the PSIC4 project, to invited presentations of other PSI validation experiments, and to a round table titled "Discussion on lessons learned from PSI validation experiments". The discussion of the round table has been based on a list of Open Questions prepared by the Validation Group of the PSIC4 project.

Since not all the Open Questions were discussed during the round table, the same questions have been later distributed to the participants, asking for their contribution. **This documents contains the outcomes of the round table and the collected contributions**. The questions are organized in two main categories:

- Questions related to the PSIC4 results,
- More general questions related to the PSI techniques.

Furthermore, the answer to the questions are classified according to their sources: ROUND TABLE (all), TRE (Alessandro Ferretti), IREA (Eugenio Sansosti), DLR (Nico Adam), TUDELFT (Ramon Hanssen), GAMMA (Tazio Strozzi). A separate section contains the INGV (Sven Borgstrom and Gianni Ricciardi) contribution.

Questions related to the PSIC4 results

- **The density of the spatial sampling provided by the 8 PSI teams over the Gardanne test site is very different:** ***Why there are large differences in the PS densities?***

  ROUND TABLE:

  - Different decisions taken by the processing teams during the data processing.
  - There is no unique definition for a PS (mostly due to different coherence definition and thresholds).
  - PSI outputs can be tailored to the specific needs of users. Different philosophies in the PS selection can be made, giving priority to one of these two goals: reliability of time-series vs. density of points. This is a key advantage of PSI, but the users have to be informed on the quality of the delivered PS. For this purpose, metadata are needed to explain limits and advantages of the PS at hand.

  TRE: In the project a common coherence threshold was not defined.

  IREA: Selection of PS is made by setting a threshold on the coherence. Different teams have used different thresholds.

  TUDELFT: Different threshold on the time coherence were used. By setting a high threshold one decides to show only the highly trustworthy data, even if the consequence is that some deformation phenomena will not be mapped. Other reasons:

- Different amount of images used.
- Different time span.
- High Doppler images used/not used.
- Difference in APS estimation.
- DEM removed/not removed.

DLR:

- The teams used different thresholds for the SCR (Signal-to-Clutter Ratio) in the course of the PS-detection. A low SCR-threshold allows a higher PS density but on the risk of mis-estimation due to the allowed higher phase noise in the observations.
- The generation of few simple differential interferograms provides a very good solution to this particular monitoring problem because the subsidence took place in a very short time interval. This is the reason that over 80% of the motion can be covered by interferometric pairs which span a time range of only three months allowing coherence even on distributed scatterers. Of course, this result is not based on permanent scatterers but suggest a very high PS-density. Some teams may have delivered such a result.

- **In particular, the PS density is in general very low in the deformation areas, even where the deformation takes place in urban areas: *Why PSI teams do not provide deformation measurements in most of the deformation areas?***

  ROUND TABLE*:*

  - Unwrapping errors (problem of sampling).
  - Strong deformation (for C-band) and non-linear evolution in time.
  - Note that by relaxing the PS selection more (less reliable) measurements could be provided

  TRE: Given the context of the PSIC4 project, where the teams have worked under "blind conditions", with no information on what is the deformation signal of interest, nobody has done an in depth analysis tailored to specific characteristics of the Gardanne mining area. A standard PSI processing has been used.

  IREA: The PSIC4 test area has shown very fast deformation that may cause loss of coherence.

  TUDELFT: Just no coherent scatterers in area. Here we mean that there are no/limited scatterers that are coherent over the entire interval/for all data. They might be coherent within a subset of acquisitions. Deformation (strongly) non-linear, while the processing started with a null-hypothesis of linear deformation. This requires higher level processing, which was not available at the time of the project. Currently, such deformations would be easier to detect.

  DLR:

  - PSI relies on stable scatters which are given by chance. Usually the scatterers are man made structures made of metal. In particular this deformation area is not covered with such structures. A test site in an urban area would provide a much better spatial sampling of the motion effect on ground. The generation of a SCR-map would show the available PS in this area (this is part of the proposed algorithmic validation).

- The estimation provides the goodness of fit parameter by the temporal coherence. In case the used displacement model is not suitable the low coherence estimates are incorrectly removed from the result assuming the detected scatterer is after all not stable in time.

- ***Why are most of the deformations missed?***

  TRE: In the mining area there are typically problems related to phase unwrapping.

  IREA: Because it is either too fast or in vegetated areas, or both.

  DLR: The displacement monitoring can be decomposed into two tasks: firstly the displacement area detection and secondly the displacement velocity estimation. I personally think that most of the teams have solved the displacement area detection problem. The shape of the displacement area looks very similar for the different teams. The reason is that the shape is not affected by the systematic underestimation of the displacement.

  GAMMA: For the current project and in comparison to other typical commercial and pre-commercial activities of GAMMA, the contact to the user was missing. This resulted on a sort of "blind" processing with unclear priorities regarding areas and time periods of major interest. For instance, for the current processing we did not focus our attention to a quantitavie analysis of the areas of higher deformation rates (partly feasible with single short-baseline interferograms) but to the slow moving regions where the potential of IPTA is higher.

- **Over the test site the deformation velocity maps of the 8 teams show discrepancies up to few millimetres. This is in contrast with the sub-millimetric values of precision that are theoretically expected from the velocity maps:** ***Considering the (expected theoretical) precision of the deformation velocity, how can we explain the biases in the velocity values?***

  ROUND TABLE: The spatial distribution and pattern of the velocity maps are needed to understand the discrepancy.

  IREA: Do not expect an accuracy better than 1 mm/year, as also demonstrated by our experiments published in the literature.

  TUDELFT:

  - Unwrapping errors.
  - Difference in APS estimation techniques (separation from non-linear deformation).
  - Different images used.
  - Difference between theoretical variance and *a posteriori* variance (idealization precision).
  - Sub-pixel position in azimuth direction taken/not taken into account.

  DLR:

  - The theoretical precision is derived for optimal conditions i.e. a linear model describes the displacement effect and the PS density allows a perfect compensation of the APS. These conditions are not given in this particular test site. Furthermore, the displacement estimation is a frequency estimation problem. I.e. the observation time (in this case the time span of the displacement which is few months only) fixes the estimation accuracy. The longer the time

span the better is the accuracy. Millimetre accuracy requires observation time spans of several years.

- The temporal sampling of 35 days and the wavelength in C-band for the current sensors make it very difficult to resolve the phase ambiguities of the observed displacement signal. Prior information would be necessary in order to adapt the motion model for the displacement estimation. The monitoring result of this test site would be better using an L-Band radar and a repeat cycle of a few days only.

GAMMA: We agree with other teams that the theoretical precision is derived for optimal conditions and that one cannot expect sub-millimetric values of precision in the specific case under discussion. In addition, the processing teams followed different processing strategies with different priorities. In order to fully understand the discrepancy the spatial distribution and pattern of the velocity maps are needed, at least for the teams that applied a PSI standard processing. For a better comparison of results we propose to concentrate on the mean velocity for the time period 1995-2000, when a regular set of ERS SAR images is available.

- **Deformation time series: the deformation time series estimated by the 8 PSI teams show in general big discrepancies with the reference time series, which are based on levelling data:** ***How can we explain the validation results of the time series?***

  IREA: A more detailed error analysis is needed to answer to this question.

  TUDELFT:

  - Errors in PSI (mainly unwrapping errors due to insufficient temporal and spatial sampling of deformation signal).
  - Different deformation regimes measured (objects with levelling benchmarks subside faster than ground (double bounce PS)).
  - Bias in the applied Kriging method.
  - Atmosphere estimation (separation from non-linear deformation).

  DLR: The processing teams followed different strategies causing different results. Some teams applied a PSI standard processing showing the correct output for a test site with no priori information. Others tuned the processing including standard D-InSAR which improves the result on this particular test site. Of course they solved the monitoring task best. But subject of the project was a cross comparison of the implemented PSI systems.

  TRE: comparing PSI data and optical levelling data is not a trivial task to perform. In our opinion a more in depth analysis should be performed.

  GAMMA: For the area characterized by the major deformation this is in particular the result of unwrapping errors due to insufficient temporal and spatial sampling of deformation signal (see also the answers to the first two questions). For “slow” moving areas structures with levelling benchmarks may subside faster than ground.

- ***Why PSI tends to underestimate the deformation?***

  ROUND TABLE:

  - Unwrapping errors, propagation of error from nearby quickly/nonlinearly moving area.
  - The question is unanswerable without access to the validation data.

DLR: The main displacement in the test site occurred in a very short time period compared to the provided data time range. By the applied unsuitable displacement models often covering the full time range the displacement is systematically spread over the full time range and thus underestimated.

TRE: unwrapping is still "the" problem of any InSAR application.

GAMMA: Because most of the SAR data are acquired for the slow moving time period and the phase ambiguity reduces the outlines of the SAR acquisitions within the very short time period of strong displacement within intervals of $\pm\pi$ (i.e. ±1.4 cm in the line-of-sight direction), instead of several centimeters.

- ***Why the start of the deformation is often seen too late?***

  IREA: Probably the PSI technique missed the time stepwise behaviour of the deformation because of phase unwrapping errors, thus underestimating the deformation rate. Moreover, the (time) high frequency applied filtering for APS separation may cut also high frequency deformation signal components.

  GAMMA: We couldn't found evidence of this in our Validation Report.

- ***Why there are big discrepancies between the time series of different teams, even in stable areas?***

  IG: Note that this issue is addressed in the Appendix of this document.

  IREA: Different teams have used different atmospheric filtering, so we should first check if some of the deformation has been confused with the APS.

  TUDELFT:

  - Different approaches: based on spatial coherence estimation or temporal coherence estimation.
  - Unwrapping errors.
  - Difference in APS estimation technique (separation from non-linear deformation).
  - Sub-pixel position in azimuth direction taken/not taken into account.

  DLR:

  - The teams use different algorithms and implementations for the co-registration. Even small co-registration errors propagate into the final estimates due to systematic phase effects in case of point scatterers and random effects in case of distributed scatterers.
  - The teams applied often linear motion models. Depending on the APS processing the non-linear motion leaked differently into the phase residuals. A spatial-temporal filtering then propagates errors into the time series explaining the deviations in the estimated displacement and the APS.

  TRE: We are quite convinced that in general it will be possible to create evidence of the good agreement among the different PS results in areas affected by moderate deformation rates, although PS densities are different due to the different thresholds used by the teams.

  GAMMA: The processing teams followed different processing strategies with different priorities. Even those teams that applied a PSI standard processing may have adopted different approaches for the estimation of the time series. In addition, phase unwrapping errors may also explain some of the inconsistencies. In any case,

we have the impression that the whole discussion regarding discrepancies between teams is biased by the results of one or two teams that adopted a different differential SAR interferometric approach.

- ***The time series represent an advanced PSI product: Can be delivered everywhere? When can be delivered?***

  IREA: Time series can be delivered in all the coherent areas provided that the user understands the significance of the coherence measurement that should be attached to it.

  TUDELFT: Yes, time series can always be delivered for every PS, it should not be considered as an advanced product, but as an intermediate product.

  DLR: From the results of this study the generation algorithm of time series should be harmonized or, at least explained in detail whenever given to customers.

  TRE: "meta-information" and in general "auxiliary information" are key-elements of any delivery of PS data. Users should be aware of the limitations of this technology as well as its potential.

  GAMMA: Yes, time series can support the interpretation of the results. For instance, strong accelerations after a relatively long period of slow displacement, even if not correctly measured, are often evident in the noise of the time series.

- ***What are reasonable values for the precision/accuracy of time series?***

  IREA: IREA experience shows an accuracy not better than 2 mm in good areas, up to 1 cm in more critical areas.

  TUDELFT:

  - Ill-posed question: precision/accuracy of which parameter?
  - Accuracy is related to true (but unknown) deformation phenomena. Assuming levelling data is deterministic ground truth, accuracy of the technique in terms of linear deformation parameter can be very bad: several cm/y.
  - Precision is related to estimated parameter, this is not necessarily relevant to phenomena of interest (consider, e.g. autonomous behaviour of points).
  - *A posteriori* precision (standard deviation) of linear velocity estimate can be better than 1 mm/y. This depends on the length (extent) of the time series.
  - Validation experiments show a standard deviation of about 2 mm for single (adjusted) measurements of time series.
  - The accuracy of the time series is also determined by the internal and external reliability of the PS-InSAR estimation procedure itself, which is related to the detectability of model imperfections and its influence on the estimated parameters (APS estimation methodology, unwrapping strategy, PS localization).
  - The stochasticity of the ambiguities should be taken into account, e.g. by a multi-modal Probability Density Function. Therefore, if the success rate of the ambiguity resolution (phase unwrapping) is not equal to one, describing the precision by standard deviations is strictly not correct.

  DLR and TRE: There is no constant precision number. The precision depends on:

  - The PS density, which affects the APS estimation and the spatial sampling of the displacement effect.

- The sensor repeat cycle which affects the temporal sampling frequency of the displacement effect.
- The number of observations and their distribution and data gaps in time and baseline.
- The used algorithm and the software implementation (critical are e.g. co-registration, resampling, PS detection, unwrapping - network solution).
- The SCR of the detected PS.
- The weather conditions.
- The reference point.

In fortunate cases 1 mm displacement per year and 1m in height can be detected and monitored.

GAMMA: From the residual phase level a theoretical precision of 2 to 3 mm is claimed for time series in the case of high point density and smooth spatial and temporal characteristics of the deformation process. However, with phase unwrapping errors as a result of relatively large displacement and inadequate temporal sampling these values are clearly not realistic.

- **Atmospheric phase screening (APS): the APS estimated by the 8 PSI teams show important differences:** ***Why are the APS between teams so different?***

  IREA: Different selection of space-time filtering thresholds.

  TUDELFT:
  - Difference in APS estimation/filtering.
  - Unwrapping errors.
  - Different orbit parameters used.
  - Orbit errors removed a priori/not removed.

  DLR: Unsuitable motion models (or temporal filters) result in leakage of non-linear motion into the residual phase which is further processed into the APS. The residual phase is affected by noise and a spatial filter usually separates the APS and the noise. Different filter settings based on assumptions of the spatial correlation of the APS and the noise level result in different APS outputs.

  TRE: different teams use different algorithms for APS estimation and different statistical characterization of the deformation signal.

- ***What is the meaning of APS?***

  IREA: It is the result of the application of a stochastic filtering on a space-time signal. If part of the signal or other possible disturbances (such as residual orbital ramps) have the same spectral characteristics assumed to belong to the APS, they will be included in the estimated APS.

  TUDELFT: Atmosphere + orbit.

  TRE: The APS should be just a spatially smoothed version of the phase residues of all differential interferograms after removing topographic and deformation components.

  GAMMA: The IPTA final results consist of heights, linear deformation rates, atmospheric phase, and non-linear displacement histories for each point target. Usual priorities for the determination of a signal related to displacement set the

estimation of the atmospheric phase at the final position. Therefore, every inaccuracy, for instance in the estimation of the orbits or in the assumption about the non-linear deformation regime, is appointed, together with the noise, to the atmosphere.

- ***What's the impact of the spatial-temporal filtering to separate non-linear deformation and atmospheric contribution?***

  IREA: Fast nonlinear deformations can be confused with APS because have the same time frequency content (fast variation with time).

  TUDELFT: Large, once part of the phase is assigned to atmosphere, it is permanently removed from the deformation profile.

  GAMMA: The discrimination of atmospheric phase, non-linear deformation, and error terms is based on their differing spatial and temporal dependencies. The atmospheric path delay is spatially correlated but assumed to be independent from pass to pass. The non-linear deformation is assumed in general to be spatially high-pass and temporally low-pass. The phase noise is random in both spatial and temporal dimensions. An entirely correct separation of the atmospheric phase term, the non-linear deformation, and phase noise is not possible, as the assumptions used in the separation are not fully valid. When the assumptions about non-linear deformation and atmospheric contribution are not valid (e.g. in the case of temporal correlated atmospheric contribution or very fast deformation) the two signals are not correctly separated.

More general questions related to the PSI techniques

- **Over the PSIC4 test site of Gardanne the PSI techniques clearly do not succeed in areas with "strong" deformations:** ***What are the limits of PSI in terms of deformation magnitude and rates?***

  IREA: The main limitation is due to phase unwrapping problems.

  TUDELFT:
  - Completely dependent on solution space boundaries (larger boundaries will induce more outliers).
  - Non-linearity is more problematic.
  - Continuity of time series is very important: temporal gaps in combination with fast localized deformation can lead to ambiguous results.
  - Dependent on wavelength.

  DLR: The actual observation limit is fixed by the currently used sensors:
  - a radar wavelength of 5.6 cm,
  - a repeat cycle of 35 days,
  - a spatial ground resolution of 10 by 30 m,
  - a look angle of about 23 degree.

  These constraints result in the following technical limits:
  - ambiguities of about 2.8 cm per phase cycle which need to be unwrapped by temporal models,
  - typically 100-200 PS per km^2 in urban areas,

- multiple scatterers inside of the resolution cell which need to be resolved in about 15% of the selected resolution cells.

There is no constant number for the limits of PSI in terms of deformation magnitude and rates. The actual processing and test site conditions (i.e. PS density, APS, spatial and temporal effect of the displacement, data availability, distribution of baselines) as well as the particular application (i.e. displacement detection or displacement rate estimation) influence the achievable result.

In fortunate cases – when a dense spatial distribution (number) of PS is given and the spatial variability of motion is smooth - 1 mm displacement per year and 1 m in height can be detected and monitored. With the upcoming sensor TERRASAR-X these limits will be overcome. The ability to monitor "strong" displacement is limited mainly by the repeat cycle and the displacement ambiguity of 2.8 cm per phase cycle. Due to the irregular sampling in time there is no constant maximum rate. The subsidence effect needs to be linear (i.e. invariable) over a long time and sufficiently sampled in time and space. With a parametric frequency estimator the upper limits are about 5 meters subsidence per year in case many Tandem interferograms can be generated and about 14 cm per year in case the expected time separation is 35 days for the sensor ERS.

GAMMA: We clearly expected this, as for this processing we did not focus our attention to the analysis of the areas of higher deformation rates (partly feasible with single short-baseline interferograms) but to the slow moving regions where the potential of IPTA is higher. Our experience, partly related to projects supported by ESRIN, suggested that ERS-ASAR IPTA performs well in the case of mining induced subsidence for urban areas with deformation rates < 2 cm/year, but that problems exist for high spatial gradients and movements that are highly non-uniform in time.

- **In the PSIC4 project the teams have worked under "blind conditions":** ***What a priori information is needed to successfully apply PSI on areas with "difficult" deformation regimes, like the mining sites?***

IREA: Any information that allows understanding the deformation trends and trend changes.

TUDELFT: No generic answer possible, depends on situation. In general, all extra reliable information on the problem at hand will create better results. Often, start time of production would be helpful.

DLR:

- Due to the displacement ambiguity of 2.8 cm per phase cycle and a sensor repeat cycle of 35 days the displacement model and initial parameters should be known. Additionally, stable areas close to the subsidence are required for the absolute calibration of the data.
- The monitoring assignment needs to be described exactly.

GAMMA: The most straightforward application of IPTA is the monitoring of slow and temporally uniform deformation. It was clear already before the PSIC4 experiment that a standard PSI processing would not perform on areas with "difficult" deformation regimes, like the mining sites. Information about the time period of underground mining would permit to analyze only SAR data acquired when displacement is small.

- **The majority of the PSI teams have used a linear model to describe the time evolution of the deformation:** ***Why only/mainly linear models are used?***

  IREA: The SBAS approach of IREA does not use a linear model, neither any other model.

  TUDELFT:

  - The SOW of the PSIC4 project stated linear velocity as deliverable.
  - Most likely null hypothesis, due to maximal redundancy of the model. Hence, parameters are best estimated.

  DLR: Due to strong atmospheric distortions some sort of model is required. With only few observations given, only low order (linear) motion models can be applied. If more observations in the temporal scale of the motion effect are available, then higher order models can be applied. In general, and as a lesson learnt, linear models should only be used over 10+ year periods, if slow effects are assumed which may often be true for geophysical effects. If sudden events are assumed, higher order models or a pre-screening process should be applied. Nevertheless, it is a reasonable displacement model in case nothing is known about the test site.

  GAMMA: In IPTA a linear model only supports phase unwrapping.

- **Some of the PSI teams have chosen the deformation reference point very far from the Gardanne area, i.e. the deformation area of interest:** ***What is the impact of using the reference point far from the area of interest?***

  IREA: The accuracy of the deformation measurement depends on the distance from the reference point, the closer the better. Choosing the reference point far from the deformation area may impair the result.

  TUDELFT:

  - In principle no impact, only unwrapping errors can influence the results.
  - Relative precision intrinsically stays the same.
  - If the reference point is closer to area of interest then impact of unwrapping errors is reduced.

  DLR: The reference point fixes the absolute displacement and height estimation. The more accurate this point is estimated the smaller is the global bias. Using only one reference point and a single observation the overall offset accuracy can only be in the order of the accuracy of this single estimation (1 mm displacement per year and 1m in height in fortunate cases). Averaging some PSs in a zero displacement area the standard deviation of the global offset error reduces by the square root of the averaged PSs count. The averaged PSs should be located closely together which guaranties that the APS is correctly compensated and the local relative displacement is right. On this particular test site, the suitable zero displacement area is located far away from the (a priori unknown) area of interest. Of course the error increases with the distance from the reference point. But this error propagation can be reduced by redundancy in the reference network. The advantage of a small global bias due to the averaging often compensates the error propagation depending on distance from the reference point.

  GAMMA: No impact in IPTA, as far as the reference point is selected in an area with high point density and with low phase noise.

- **The PSIC4 results show important limitations of the PSI techniques in areas of strong deformation, and in some cases in areas with moderate deformation:** ***Are the PSIC4 results the best we can get of PSI at this moment?***

  IREA: For the type of test site and without a priori information, yes. It is very likely that all the teams worked at their best.

  TUDELFT: No, algorithms are strongly improved over the years.

  GAMMA: The PSIC4 exercise was intended to be a standard PSI processing.

  DLR: Please see answer on "What are the limits of PSI in terms of deformation magnitude and rates?". The unfortunate PSIC4 results are specific to the current sensors and this particular test site. These results should not be generalised.

- ***What are the improvement we can wait for the techniques?***

  TUDELFT:

  - Improved unwrapping (3D).
  - Improved APS estimation.
  - Improved stochastic model.
  - Improved co-registration.
  - Sub-pixel position taken into account.

  GAMMA: In addition to the other improvements described by other teams we can mention the separate estimation of the point heights and deformation histories.

- **The teams have worked under "blind conditions", with no a priori information on deformation type, driving mechanism, deformation magnitude, temporal evolution, and on the deformation signal of interest:** ***Is there a need for a "less blind" PSI validation experiment?***

  IREA: It is advisable to more deeply analyse the result of this experiment before setting up a new one.

  TUDELFT: Not directly, because of risk in overselling. First, improvements need to be made on existing examples. Once algorithms have evolved, there is a need for another validation experiment. This should not be less blind! In a less blind experiment, the final results will be significantly biased by the a priori information, which makes an objective evaluation of the technique more difficult.

  DLR: I prefer an algorithmic validation which is complementary to the PSIC4 geo-validation.

  GAMMA: First of all, we think that a proper conclusion of the PSIC4 experiment is necessary. This include e.g. putting the whole experiment in the right context: on one side, we had processing teams that mostly applied a standard technique tailored for the monitoring of slow and temporally uniform deformation, and, on the other side, a validation team that concentrated the efforts over the mining area, where limitations of the techniques were well known before the experiment.

Additional comments from INGV

The observations below come from the writers experience and are related to the particular conditions we have in the Neapolitan Volcanic District, that are:

- The availability of a large geodetic monitoring system, with permanent and periodically measured networks.

- The low deformation rates of the Phlegrean Fields area.

We likely suppose that, with different conditions, something could change.

- **The advantages of DInSAR in comparison with classical geodetic techniques**
  - A good space coverage

    Ground networks, also allowing very high accuracies, give an information related only to a certain number of measuring points: this limitation can be greatly relieved by exploiting DInSAR which allows to retrieve an information on wide areas, also indicating eventual migrations of the deformation field in the SAR scene to be afterwards investigated with classical techniques. In general, SAR results can provide Volcanological Observatories/Civil Protection Agencies with a clear indication about suitable locations for geodetic networks installation.
  - A good time coverage

    Field measurements over the geodetic networks (periodic networks) are typically carried out once or twice a year (unless in case of crisis, when data sampling is strongly increased). Therefore the good time coverage (35 days/swath for the ENVISAT ASAR sensor) of SAR data vs. the mean repetition time of field measurements represents a strong improvement for monitoring purposes. In particular, just in the case of PSI, the possibility to get (long) ground deformation time-series over coherent targets has allowed in some cases to retrieve the information not available from classical techniques in the time interval between the field measurements. This is not true when dealing with permanent networks (i.e. GPS) but also in this case we must consider the poor space coverage of permanent stations on the area of interest, mainly related to economical reasons, besides some technical and logistic ones.
  - A (drastically) lower cost

    The great advantage to get an information over an area of 100 by 100 km^2 at a very low cost (25 €/scene in CAT-1 regime) shall be mentioned, especially when compared with the cost of each geodetic field measurement, in the order of some to many thousands €. Also considering some drawbacks in the SAR scene due, for instance, to low or no coherence in some areas, the huge amount of information we can get for that price has no comparison with the (poor, also if very precise) results from classical measurements.
  - An optimization of the monitoring system

    The joint use of SAR and geodetic data represents an optimization of the monitoring system for Volcanological Observatories/Civil Protection Agencies which allows primarily to get more information about the deformation field, at both large and small scale. Secondly, the availability of SAR and geodetic data allows the choice of the geodetic surveying techniques more suited as a function of the deformation rate (i.e. the volcanic risk level) of the area: typically classical ground techniques shall be used with no or low dynamics, whereas geodetic surveying from spaceborne sensors shall be recommended during pre-eruptive/eruptive phases, when the access to the volcano is not possible and/or the traditional monitoring networks are not available due to blackout.

- **More observations on the comparison among DInSAR and classical geodetic techniques**

In the last years we made some comparisons between SAR data (processed by IREA-CNR) and GPS data from our permanent network and, more recently, with levelling data from our levelling network. As a general comment we can say that the comparative analysis between different data sets has shown a good agreement between the different techniques in both space and time, as also reported by some papers on international journals published in these years. The main limitations towards a more precise comparison are due to the poor space coverage of the permanent stations (i.e. GPS) and the poor time coverage of geodetic data from periodic networks (i.e. levelling) in comparison with the good space-time coverage of SAR data. The main issue is the different acquisition geometry, i.e. into the Radar Line of Sight (LOS) for SAR, the 3D displacement vector from GPS and the only vertical displacement from levelling. However, this is not critical because we can get:

- In the first case, the GPS data re-projected into the radar LOS (depending on the availability of a 3D data), as it was done during the 2000 uplift event in the Phlegrean Fields, with about 5 cm of deformation at that time.
- In the second case, a quantitative comparison of levelling data with SAR results, taking into account the possibility to split SAR data, when they are available from both ascending and descending tracks, into the vertical (comparable with levelling) and the EW components of ground motion.

In our opinion, we think SAR data should not be compared with GPS data, because they both suffer from the same drawback, that is the propagation of an electromagnetic pulse through the atmosphere, well known by the SAR experts as the Atmospheric Phase Screen (APS). In order to precisely validate/compare SAR data, we need to exploit a classical technique which does not present such a drawback, which is in this case the levelling technique.

Levelling has for sure other drawbacks in comparison with SAR (i.e. time/money consuming, etc.) but no the atmosphere crossing of the signal as in SAR. On the other hand, in our experience we can also state that SAR Interferometry cannot be the only monitoring technique, because of some limits in comparison with the classical ones. As an example, since the last two years we are recording a very small uplift in the Phlegrean Fields (about 3 cm) which was pointed out from the permanent networks (tilmetric and GPS) but, at the beginning, not from SAR, due to the low deformation rate. Only after a certain time SAR data gave interesting indications, when the deformation signal was strong enough to be detected by SAR (need for X band data?).

Another critical issue is the choice of the reference point for SAR data processing, which could be given by ground networks; in fact, these are conceived to have an external point of the network, outside the maximum deformation area, to be used as a reference for differential processing. It was successfully done during the 2000 uplift event, when the SAR reference point has been the same of GPS.

PSIC4

APPENDIX B
of the Final Report of PSIC4:

Intercomparison Provence

Marta Agudo, Michele Crosetto
Institute of Geomatics

Final version
Castelldefels, 15 March 2007

INSTITUT DE GEOMÀTICA

GMES TERRAFIRMA

INDEX

1. INTRODUCTION

Within the PSIC4 (Persistent Scatterers Interferometry Codes Cross Comparison and Certification for Long term differential interferometry) project, eight different teams specialized in Persistent Scatterers Interferometry (PSI) processed a large stack of ERS and Envisat SAR images over the same test site. The site was located around the mining area of Gardanne, in Provence, in southern France. Though the areas covered by the different PSI teams are considerably larger than the Gardanne mining area, see Figure 1, the PSIC4 project was mainly focused on this area, where most of the ground deformations occur and the ground truths, coming from repeated levelling campaigns, are available.

This document describes the main results of the inter-comparison of the products of four PSI teams involved in the PSIC4 project. This inter-comparison, which is named "Provence Inter-comparison", has been performed within the Terrafirma project. For this reason, the analysed data correspond to four of the five Operational Service Providers (OSPs) of the Terrafirma project, namely Altamira Information, DLR, Gamma Remote Sensing and TeleRilevamento Europa (TRE). The analysis described in this document was started after the PSIC4 Final Presentation and the PSI validation workshop, held in ESRIN the 18th and 19th September 2006, where the final results of the PSIC4 project were presented. The main motivation for starting this complementary study was to take advantage of the full PSIC4 dataset by extending the analysis beyond the Gardanne mining area, thus getting new insights into this unique set of PS data. It is worth mentioning that an important additional motivation to perform this study was to collect inputs for the preparation of a new validation experiment, to be carried out in 2007 as part of the Terrafirma project. The main focus of this study is on the global inter-comparison performances of four TF OSPs. The inter-comparison is expected to answer this key question: how consistent are the results between different teams? It is worth emphasizing that all the results discussed in this document refer to differences between PSI teams. These results should not be confused with the errors estimated in a validation experiment by comparing PS results against ground truths.

Even though PSIC4 and the Provence inter-comparison share the same PS dataset, there are a few key differences between the two projects, which are briefly listed below.

1) The PSIC4 dataset includes both geocoded PS data and data in the original radar coordinates. The validation and inter-comparison results of PSIC4 are based on the geocoded PS data. By contrast, the Provence inter-comparison is exclusively based on the original PS data, i.e. those defined in the radar image space. The analysis in the radar image space is performed at full resolution of the SAR SLC images. This inter-comparison has the advantage of minimizing the "validation errors". In fact, it avoids the errors associated with the pre-processing steps needed to validate and inter-compare the geocoded PS data (e.g. correct for the geocoding errors, spatial interpolation, etc.).
2) The results of the PSIC4 project refer to the Gardanne mining area, which is about 100 km^2. This is the area where the bigger land deformations occurs. By contrast, the Provence inter-comparison concerns the entire areas covered by the PS data, which mainly include stable or moderate deforming areas. This corresponds to at least 925 km^2, see Figure 1.
3) The PSIC4 project involved both validation activities against levelling data and inter-comparison activities. The Provence inter-comparison only involves inter-comparison of the deformation products between the four PSI teams.

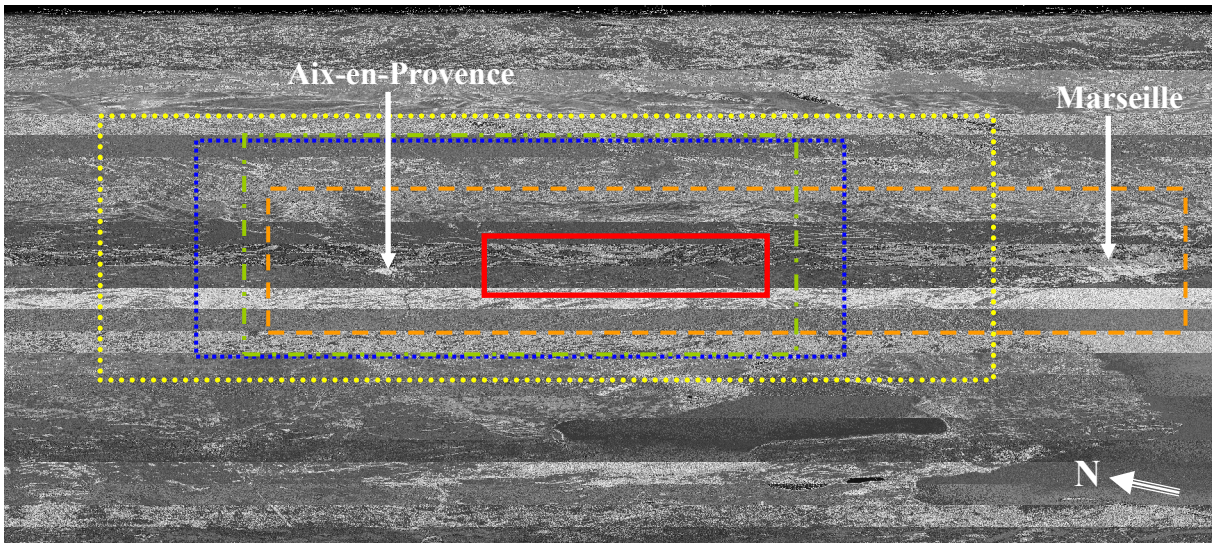


Figure 1. Mean SAR amplitude over the studied area. The red frame indicates the approximate mining area of Gardanne analysed in PSIC4. The other frames show the PS coverage of the four teams considered in the Provence inter-comparison.

4) The analysis performed in the PSIC4 project concerns different aspects of the PSI products, like the deformation velocity maps, time series, atmospheric phase screens, geocoding errors, etc. In the Provence inter-comparison the analysis is restricted to two key PSI products: deformation velocities, and deformation time series.

5) The PSIC4 project involved the analysis of the results of eight PSI teams. Two of them performed a coherence-based analysis, working at low resolution with respect to the original SLC image resolution. The Provence inter-comparison is restricted to four of the five OSPs of the Terrafirma project. The fifth OPS, NPA, was not involved in the PSIC4 project. In the PSIC4 project the identity of the teams was kept anonymous by simply labelling the results T1 to T8. In order to maintain anonymity in this study, its results have been arbitrarily renamed TA, TB, TC and TD.

As already mentioned above, the Provence inter-comparison is based on the PSIC4 dataset, and in particular on the original PS tables generated by the four teams. The PS radar coordinates of these tables do not refer to co-registered SAR data, i.e. the PS of different teams with the same radar coordinates do not correspond to the same footprint on the ground. Furthermore, some teams provided PS data at the original SLC resolution, i.e. with a grid that has the same size of the SAR SLC images, while others provided denser PS data, coming from over-sampled SAR data. For this reason, before starting the inter-comparison some pre-processing was performed to achieve homogeneous and co-registered datasets. This step is addressed in Section 2. The analysis of the deformation velocity maps is described in Section 3, while the study of the deformation time series is discussed in Section 4. The conclusions of this work are summarized in Section 5. Finally, an Appendix includes different complementary results of this work.

2. PRE-PROCESSING STEPS

The input data of this work were the original PS table generated in the PSIC4 project. Before starting the inter-comparison of the PS data, few pre-processing steps were performed, which are summarized in the following two tasks:

1) Homogenizing the datasets, which involved the transformation of the input PS data in the original geometry of the SLC images, and the co-registration of the different PS datasets.
2) Referring the four PS datasets to the same stable area.

The two tasks are concisely described in the following sections.

2.1 Input data

The datasets used for the Provence inter-comparison came from the PS tables generated by the teams in the PSIC4 project. Each team provided for each measured PS the following data:

- The image coordinates: range, azimuth.
- The geocoded WGS84 geographical coordinates: latitude, longitude, ellipsoidal height.
- The mean velocity over the period 1992-2000, measured in the LOS (Line-Of-Sight).
- The time series of LOS heights, with one value for each processed SAR acquisition date.
- The APS (Atmosphere Phase Screening) of each processed SAR image.
- The PS coherence, between 0 and 1, which indicates the quality of the PS measurements.

The four datasets considered in this work include a number of PS which ranges from about 75000 to 95000. The difference in the PS density is due to the fact the teams used different criteria during the PS processing, and in particular during the PS selection.

2.2 Homogenizing the datasets

The PS tables generated by the teams in the PSIC4 project are not co-registered, and therefore the PS data cannot be directly compared. This is due to different reasons:

- Some teams used the raw SAR data and focused them using their own SAR focusing algorithms, while others used the already focused SLC images.
- The teams used different conventions to make the transformation from the original SAR raster data to the PS tables. In fact, in this transformation one has to assign some coordinates to the first pixel of the SAR images, e.g. (0.0; 0.0), (0.5; 0.5) or (1.0; 1.0). The values used in the transformation were not documented in the PS tables.
- Finally, some teams made the PS tables starting from SAR data defined at the original resolution of the SLC images, while others started from over-sampled SAR data, thus providing “sub-pixel” locations of the PS.

In order to homogenize the datasets the following operations were performed:

1) Refer the PS data to the raster grid geometry of the original 1x1 SLC images. For the teams that used over-sampled SAR data this involved averaging the multiple PS located in the same pixel (e.g. with an over-sampling by factor two in range one may have two PS per pixel), and rounding the PS coordinates to the nearest integer.
2) Co-registration of the four 1x1 raster grids. The geometry of the SAR image with orbit number 20460 was chosen as a reference. In order to check the co-registration, a mean SAR amplitude image of the studied area was used, see Figure 1. The amplitude image

was compared with the plots of the PS, in an attempt to assess the relative shifts of the teams. However, this operation is not very accurate. The teams were contacted to assess the quality of the estimated co-registration. Finally, a further co-registration check was performed by computing the two image shifts that, for each team, maximize the number of PS coincidences between the teams. This was performed using a tool kindly provided by Nico Adam (DLR). However, for at least one team there were two shifts, separated by a difference of 1 pixel, that gave a similar number of coincidences.

It is important to underline that the co-registration realized in this study suffers an important limitation: it was carried out "a posteriori", only using the PS tables. Much better co-registration performances could have been achieved by using the teams' SLC images. This represents an **important input for future PSI validation experiments**:

Plan the experiment in order to get a very accurate dataset co-registration in the radar space, directly from the PSI teams. This requires fixing a priori:
- the reference SLC image, with a defined image resolution, e.g. the full SLC resolution;
- the type of radar coordinates, to avoid confusion it is better to use integer values;
- the convention to be used for the "coordinates of 1st pixel", e.g. 0,0; 0.5,0.5; 1,1;

If possible, require extra data, e.g. for each team a PS map overlaid with the mean SAR amplitude image. If this is not done, there is the risk of wasting resources, achieving, "a posteriori", poor accuracy in PS dataset co-registration.

2.3 Referring the PS datasets to the same stable area

Due to the relative nature of DInSAR measurements, the PS data are usually referred to a point, which is chosen as reference for all the other measurements. In order to inter-compare the PS of different teams, the datasets have to be referred to the same point. The PS tables generated by the teams in the PSIC4 project refer to different reference points, see Figure 2. This results in small global shifts between their velocity maps, which in turn cause differences in the time series that are linear trends. An example of such shifts is shown in Section A1.1 in the Appendix.

To avoid the above mentioned shifts and trends, we decided to refer the four PS datasets to the same reference area, which was chosen in a region where the PS of all four teams show no deformation. The approximate location of this stable area is shown in Figure 2 by a yellow frame, while the zoom over this area is shown in Figure 3. Over a small portion of this area, indicated by a red square, we computed the statistics for the velocities of the four teams, see the table in Figure 3. Having checked that in the chosen area there are not PS showing autonomous movements, we used the estimated mean velocity values to refer all the PS tables to the same reference area. This was done by applying a shift to the PS velocities, and removing the corresponding linear trend from the time series.

It is worth noting that mean velocity values shown in the table in Figure 3 are similar to those estimated in the PSIC4 project for the geocoded PS. Apart from a global shift of about 0.5 mm/yr, due to the fact that different reference areas were used, the relative differences between the teams, e.g. $mean_{TA}$ - $mean_{TB}$, are very close to the values assesses in PSIC4. In other words, the values estimated in this study are basically consistent with those of PSIC4.

From this task there is another **input for future PSI validation experiments**:

Define the stable reference area a priori and require the PS tables to be referenced to this area. Note that a second reference area is needed for the PS topographical errors.

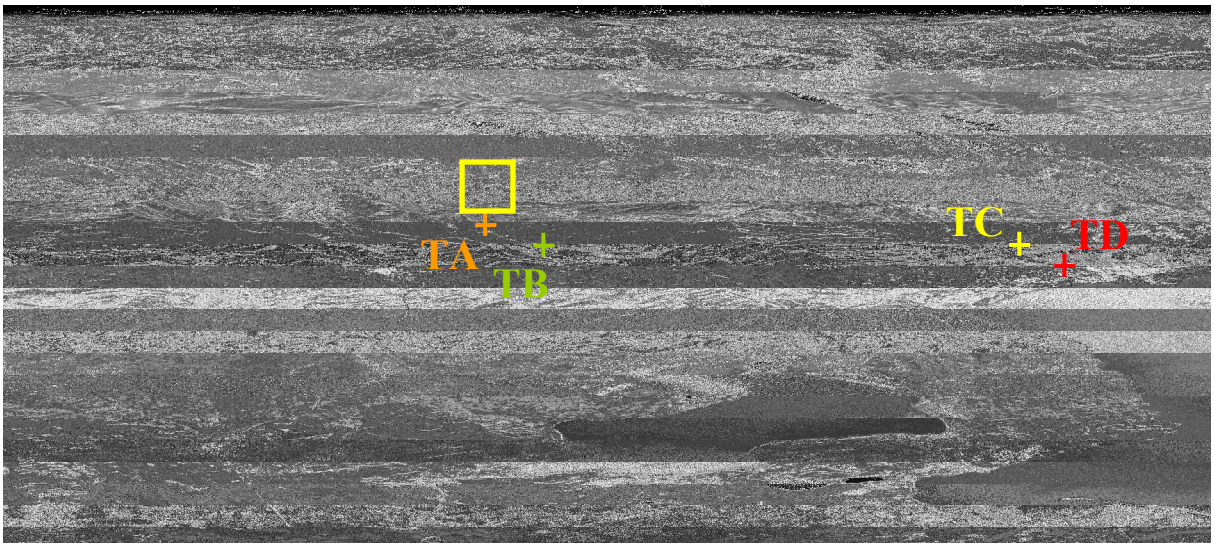
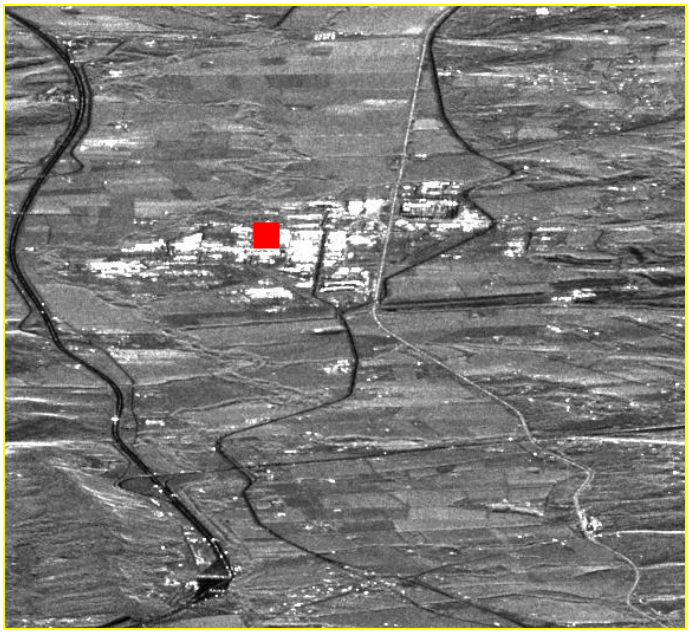


Figure 2. Approximate location of the reference points used in PSIC4 by the four teams. The yellow frame indicates the approximate location of this stable area used in this work.



[mm/yr]	Mean	Stdev	# PS
TA	1.21	0.56	27
TB	1.41	0.43	40
TC	-0.76	0.55	36
TD	-0.26	0.33	6

Figure 3. Zoom over the yellow frame from Figure 2. The statistics of the PS velocities of the four teams, shown in the table, were computed over the small area shown by a red square. TA and TB, which used reference points located close to each other, have similar mean values. The same occurs for TC and TD.

3. INTER-COMPARISON OF THE VELOCITY MAPS

This section describes the inter-comparison of the velocity maps generated by the four teams. The original four maps, before referring them to the same area, are shown in Section A1.2 of the Appendix. The analysis considers the six possible differences between the four teams. For each of the six pairs of teams, the velocity differences were computed on the common PS, i.e. the locations where both teams have provided the PS measurements.

The first statistics of such differences are shown in Table 1, which concern the entire common areas covered by the teams, excepting the Gardanne area. In Table 2 the same type of statistics are shown, including the mining area of Gardanne. The number of common PS per each team pair are indicated in the last column of both tables. In Table 1 they range from 7700 to 17900, while in Table 2 they range from 8900 to 27400. The statistics of these tables, which come from rather large samples of velocity differences, provide information on the global inter-comparison behaviour of the teams. It can be seen that the statistics shown in Tables 1 and 2 are rather similar. For this reason, below we mainly refer to Table 2, which corresponds to the entire common areas covered by the teams, including the Gardanne area.

The mean differences in Table 2 are included between – 0.37 and 0.18 mm/yr. Even though these values appear qualitatively small, it is worth analysing these values quantitatively. Performing for the six pairs a paired T-test on the analysed PS velocities, which coincides with a single sample T-test on the mean of the differences between the teams, with a confidence interval of 99% (α=0.01), only the TD-TC difference mean is not significantly different from zero. That is, all the other pairs have means of the differences that are significantly different from zero.

[mm/yr]	Max.	Min.	Mean	Stdev	# PS
TD-TB	17.46	-13.71	0.16	0.73	7716
TD-TA	6.65	-10.26	-0.17	0.56	5930
TD-TC	6.98	-17.04	-0.02	0.71	6560
TB-TA	13.26	-14.98	-0.34	0.60	17892
TB-TC	10.13	-14.42	-0.43	0.81	15564
TA-TC	6.63	-3.87	-0.07	0.70	7952

Table 1. Statistics for the velocity differences computed over all the common areas, except the Gardanne area.

[mm/yr]	Max.	Min.	Mean	Stdev	# PS
TD-TB	17.46	-13.71	0.18	0.70	12348
TD-TA	7.92	-11.46	-0.09	0.63	9579
TD-TC	11.24	-17.04	-0.02	0.77	8909
TB-TA	13.26	-14.98	-0.32	0.58	27445
TB-TC	10.13	-14.42	-0.37	0.80	23418
TA-TC	6.63	-11.47	-0.08	0.79	12471

Table 2. Statistics for the velocity differences computed over all the common areas, including the Gardanne mining area.

[mm/yr]	|Δv|<1	|Δv|<2	|Δv|<3	|Δv|<4	|Δv|<5
	%	%	%	%	%
TD-TB	94.45	98.74	99.43	99.73	99.83
TD-TA	94.25	99.06	99.75	99.88	99.95
TD-TC	90.79	98.99	99.65	99.77	99.85
TB-TA	93.27	99.28	99.75	99.89	99.92
TB-TC	74.60	98.48	99.79	99.88	99.90
TA-TC	86.57	98.87	99.79	99.95	99.99

Table 3. Percentage of PS differences that are, in absolute value, below a given threshold. This was computed over all the common areas, except the Gardanne area.

[mm/yr]	|Δv|<1	|Δv|<2	|Δv|<3	|Δv|<4	|Δv|<5
	%	%	%	%	%
TD-TB	94.24	98.74	99.44	99.70	99.81
TD-TA	94.12	98.80	99.49	99.70	99.85
TD-TC	89.64	98.63	99.46	99.63	99.75
TB-TA	93.59	99.17	99.72	99.89	99.93
TB-TC	76.98	98.54	99.72	99.87	99.90
TA-TC	84.64	98.40	99.49	99.76	99.87

Table 4. Percentage of PS differences that are, in absolute value, below a given threshold. This was computed over all the common areas, including the Gardanne area.

The standard deviation of the velocity differences ranges between 0.58 and 0.80 mm/yr. These values provide a rough idea on how these differences are distributed. A more complete description of the distributions is given in Tables 3 and 4, which show, in percentage over the total number, the absolute differences below 1 to 5 mm/yr. In agreement with the standard deviations, it can be seen that the great majority of the velocity differences are in module below 1 mm/yr. In the statistics that include Gardanne (Table 4), this occurs in at least 77% of the common PS, for all the pairs. The percentages globally increase for the statistics that do not include Gardanne (Table 3), even though for one pair (TB - TC) the percentage is lower, 74.6%. In Table 3 almost the totality of the velocity differences are, in absolute value, below 3 mm/yr. The number of PS with differences are above 3 mm/yr range from 17 (TA - TC) to 45 (TB - TA). In Table 4 the numbers are similar, even though the differences with larger magnitude increase. For instance, the number of PS with differences are above 3 mm/yr range from 48 (TD - TC) to 69 (TD - TB). This is an interesting result because over the mining area of Gardanne, where the PS validation gives its worst results, the quality of the PS inter-comparison does not strongly decrease.

An additional comment is needed on the min. and max. differences, shown in Tables 1 and 2. The min. values in Table 2 range from -11.46 to -17.04 mm/yr, and the max. from 6.63 to 17.46 mm/yr. These values, which are quite large if compared with the standard deviations, correspond to particular locations, where, at least one of the two compared teams has large errors in their estimated PS velocity. It is worth noting that these cases are rather rare, e.g., the PS with absolute differences above 5 mm/yr are 0.07% (TB - TA) to 0.25% (TD - TC).

[mm/yr]	TD-TB			TD-TA			TD-TC		
	Mean Δv	**Stdev Δv**	**# PS**	**Mean Δv**	**Stdev Δv**	**# PS**	**Mean Δv**	**Stdev Δv**	**# PS**
-∞≤v≤-4	-0.25	1.85	247	-0.28	2.41	173	-0.36	2.64	175
-4≤v≤-3	-0.14	1.07	179	-0.27	1.25	131	0.15	1.14	145
-3≤v≤-2	-0.02	0.78	417	-0.28	0.82	308	0.12	0.98	267
-2≤v≤-1	0.00	0.63	1547	-0.28	0.59	1220	0.02	0.73	815
-1≤v≤0	0.13	0.52	6183	-0.17	0.44	4911	-0.07	0.60	3952
0≤v≤1	0.36	0.55	3389	0.12	0.43	2565	-0.01	0.57	3281
1≤v≤2	0.71	1.07	323	0.50	0.71	226	0.40	1.27	228
2≤v≤3	1.23	1.64	41	0.91	1.14	33	1.28	1.78	22
3≤v≤4	2.10	3.07	11	0.51	0.92	5	1.38	1.42	7
4≤v≤∞	3.04	5.31	11	3.04	2.51	7	0.25	1.93	17

[mm/yr]	TB-TA			TB-TC			TA-TC		
	Mean Δv	**Stdev Δv**	**# PS**	**Mean Δv**	**Stdev Δv**	**# PS**	**Mean Δv**	**Stdev Δv**	**# PS**
-∞≤v≤-4	-0.70	2.74	326	-0.21	2.35	325	-0.76	2.53	165
-4≤v≤-3	-0.46	0.76	286	0.07	0.81	315	0.31	0.92	159
-3≤v≤-2	-0.53	0.76	799	-0.09	0.85	779	0.09	0.83	350
-2≤v≤-1	-0.47	0.56	3707	-0.24	0.77	3101	-0.03	0.78	1346
-1≤v≤0	-0.33	0.41	16661	-0.43	0.72	14619	-0.15	0.72	5701
0≤v≤1	-0.17	0.46	5276	-0.42	0.69	3974	-0.05	0.67	4365
1≤v≤2	0.26	0.79	336	-0.05	1.21	250	0.35	1.02	344
2≤v≤3	0.51	1.36	44	0.24	1.81	43	0.78	1.78	33
3≤v≤4	-0.24	0.20	6	0.13	1.90	9	-0.44	0.30	7
4≤v≤∞	6.19	6.36	4	4.45	4.44	3	-0.57	-	1

Table 5. Statistics for the differences as function of the velocity (subsidence has negative velocities). This was computed over all the common areas, including Gardanne.

It is interesting to analyse the statistics for the differences as a function of the velocity classes. In Table 5 the classification was based on the velocities of the first team of each pair, which is indicated in the first column of Table 5. As expected, the best PS differences correspond to the velocity around zero, e.g. between -1 and 1 mm/yr. (TB - TA) shows the best performances, with a standard deviation of 0.41 mm/yr (-1≤v≤0). When the velocity module increases, for both negative and positive values, the standard deviations slowly degrade. In the subsidence values, there is a noticeable decrease when the velocity is below -3 or -4 mm/yr.

Finally, to get an idea of the spatial distribution of the difference, two zooms on the maps of the differences are shown in Figure 4, 5a and 5b. For visualization purposes the PS size has been increased. Figure 4 corresponds to a portion of the area from Figure 3, located close to the reference area. There are dominant greenish colours, which correspond to difference around zero, plus some localized bigger differences, e.g. one in red for (TD – TB), of about 5 mm/yr. A wider area is shown in Figures 5a and 5b, which is included in the Gardanne mining area. Also in this case there are some locations where the PS show big discrepancies.

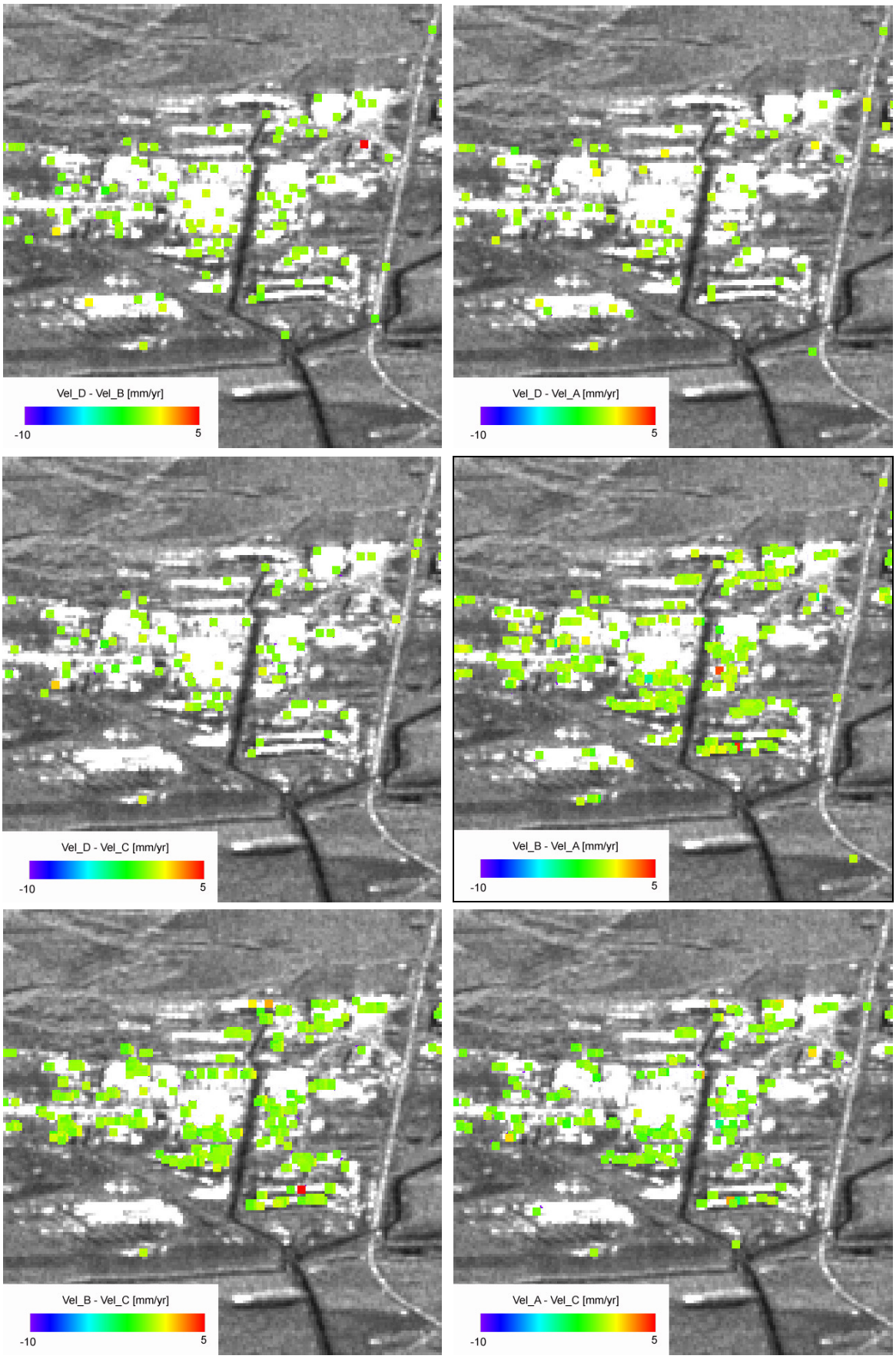


Figure 4. Zoom over the six maps of the velocity differences, superposed to the mean SAR amplitude. This zoom includes the reference area used for the PS datasets.

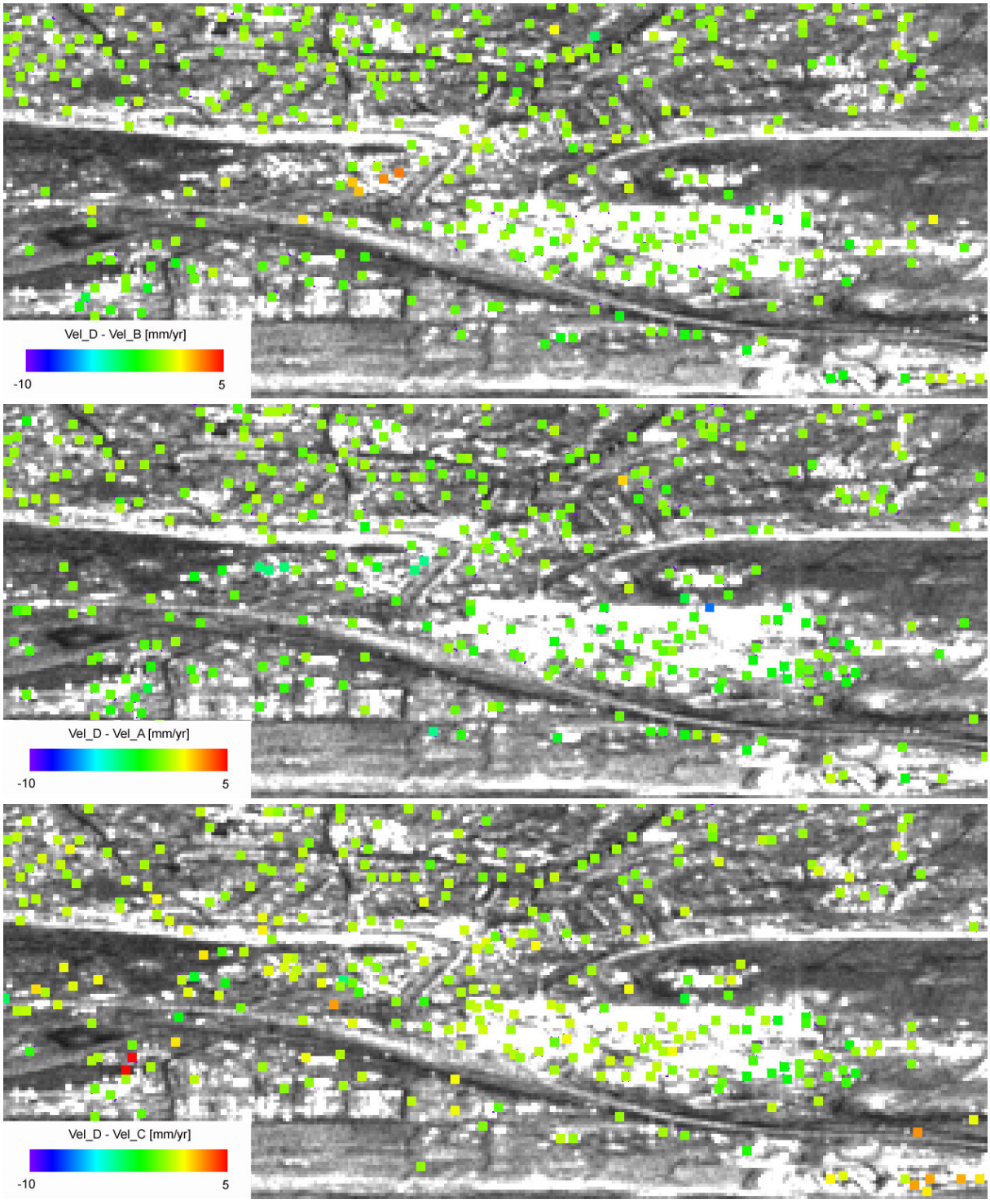


Figure 5a: Zoom over three maps of the velocity differences, superposed to the mean SAR amplitude. This zoom covers a portion of the Gardanne mining area.

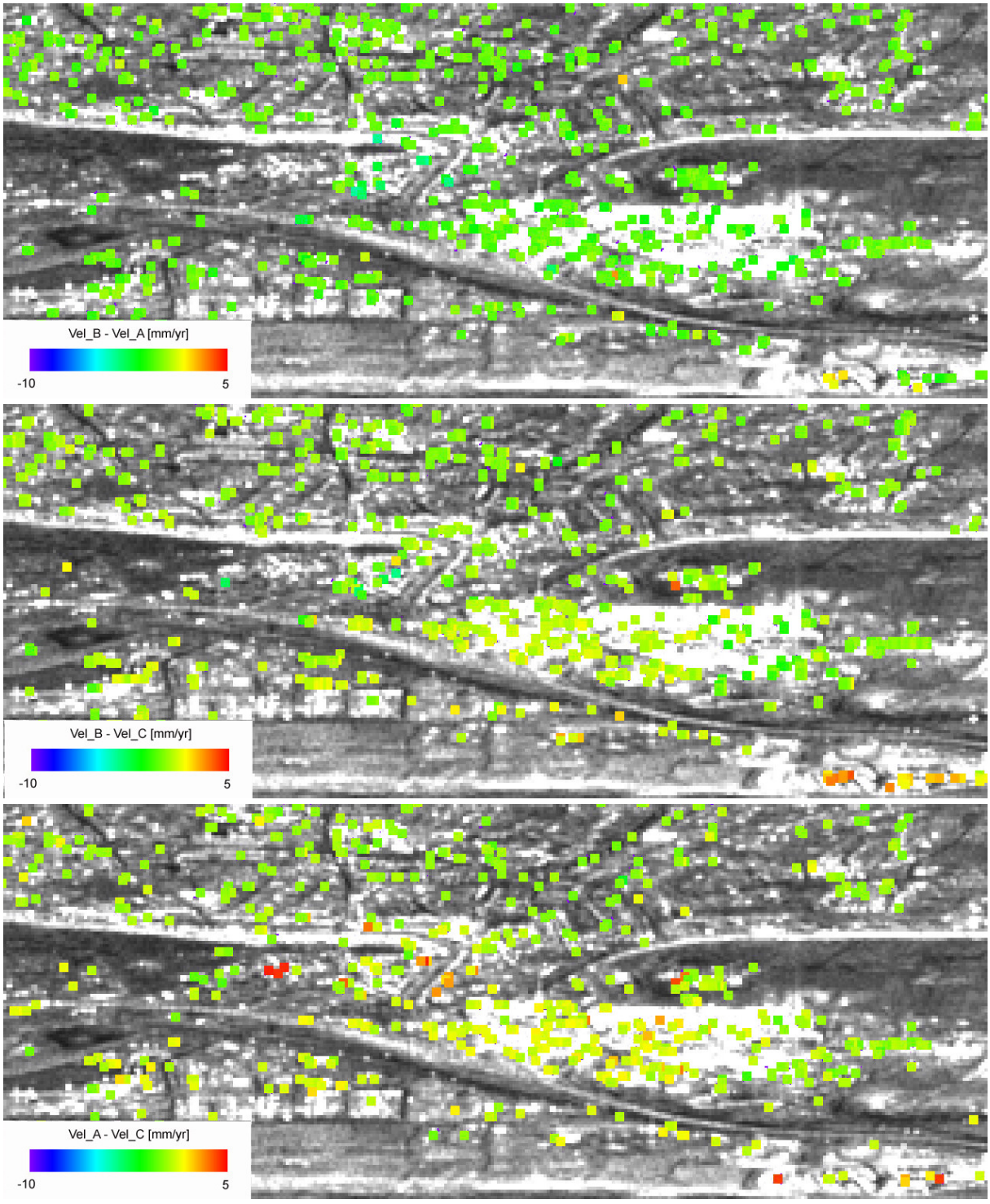


Figure 5b: Zoom over three maps of the velocity differences, superposed to the mean SAR amplitude. This zoom covers a portion of the Gardanne mining area.

3.1 Additional study: traffic light maps

Besides the standard PS products, like the deformation velocity maps and time series, simpler deformation monitoring products can be defined. One example is given by a screening, traffic light map, which indicates stable and unstable areas. Since in the Terrafirma project there is an ongoing discussion about new deformation products, an additional inter-comparison study has been performed on the velocity maps of the Provence dataset. The study concerns traffic light maps, which include three deformation classes: “stable”, “subsidence” and “upsidence”.

Four traffic light maps were computed by reclassifying the four velocity maps. Absolute velocity values below 2 mm/yr were assigned to the “stable area” class (between 92 and 95% of pixels), the velocity values below – 2 mm/yr were assigned to “subsidence” (between 4 and 7% of pixels), and those above 2 mm/yr to the “upsidence” class (less than 1.5% of pixels). The four traffic light maps, shown in Section A1.3 of the Appendix, were then inter-compared, see Table 6. From this table one may notice that for the “stable” class there are high percentages of agreement between the four traffic light maps, with values between 97.2 and 99.1%. The levels of agreement decrease for the “subsidence” (from 56.1 to 86.6%) and “upsidence” (from 11.8 to 62.1%) classes. These different percentages between the three classes are influenced by the different number of pixels belonging to each class. For instance, the “stable area” class is very populated and has the best levels of agreement between teams.

Regarding these results, it is important to emphasize that the analysis suffers a strong limitation. In fact, it does not involve a true “new PSI product”, like a traffic light map. It is simply computed on reclassified velocity maps, which were not generated with the aim of deriving traffic light maps. These maps represent a simplified PSI product with respect to the velocity maps, which however can offer other advantages, like an improved reliability or a wider coverage with respect to the velocity maps. For instance, the velocity maps over the mining area of Gardanne have poor PS density. To get a better coverage over this area a specific procedure to generate the traffic light maps should be used, which differs from the standard PSI processing used for the velocity maps. This was clearly not done in this study.

TD/TB	subs	stable	ups
subs	70.05	29.68	0.27
stable	1.90	97.70	0.40
ups	3.45	34.48	62.07

TB/TA	Subs	Stable	ups
subs	86.55	13.45	0.00
stable	1.83	98.09	0.08
ups	2.78	55.56	41.67

TD/TA	subs	stable	ups
subs	78.57	21.43	0.00
stable	2.21	97.23	0.56
ups	16.67	33.33	50.00

TB/TC	Subs	Stable	ups
Subs	62.59	36.93	0.48
stable	1.28	98.64	0.08
ups	0.87	86.09	13.04

TD/TC	subs	stable	ups
subs	70.98	29.02	0.00
stable	0.67	99.08	0.25
ups	3.13	43.75	53.13

TA/TC	Subs	Stable	ups
subs	56.14	43.86	0.00
stable	1.30	98.38	0.32
ups	0.00	88.24	11.76

Table 6. Inter-comparison of the traffic light maps. The results are organized by rows. Considering the first table, TD/TB, the 70.05% of “subsidence” pixels in TD are in the class “subsidence”, 29.68% are in the “stable area”, and 0.27% are in the “upsidence” class in TB.

4. INTER-COMPARISON OF THE TIME SERIES

This section describes the inter-comparison of the deformation time series generated by the four teams. The analysis was performed by considering the six possible differences between the teams. For each pair of teams and for each common PS, the analysis was computed on the samples of the time series where both teams have PS measurements. The main statistics of this analysis are shown in Table 7, which concern all the common areas covered by the teams, including the Gardanne area. These statistics are global, i.e. they concern all common PS between pairs of teams. They were computed as follows:

1) For each pair of teams identify all common PS;
2) For each PS identify the common samples of the two time series, and compute the difference between time series in each sample;
3) For each common PS, compute three main parameters: the mean, the standard deviation and the maximum absolute values of the above differences;
4) Considering all common PS of a given team pair, compute the mean, standard deviation, minimum and maximum values of the parameters computed in the previous step, i.e. the mean, the standard deviation, and the maximum absolute value of the differences.

In table 7 the statistics over the three main parameters are listed by columns.

TD-TB	Max abs [mm]	Mean [mm]	Stdev [mm]
max.	207.67	89.66	45.58
min.	3.63	-77.02	1.40
mean	15.20	-0.78	3.87
stdev	7.52	5.84	1.54

TD-TA	Max abs [mm]	Mean [mm]	Stdev [mm]
max.	108.03	49.60	30.25
min.	3.68	-54.51	1.45
mean	19.00	-0.11	4.34
stdev	8.00	6.54	1.53

TD-TC	Max abs [mm]	Mean [mm]	Stdev [mm]
max.	196.58	60.12	45.95
min.	4.47	-92.86	1.71
Mean	16.03	0.88	3.94
Stdev	7.51	5.84	1.45

TB-TA	Max abs [mm]	Mean [mm]	Stdev [mm]
max.	155.67	68.63	36.76
min.	3.07	-82.07	1.21
mean	13.85	0.86	3.57
stdev	5.90	4.99	1.21

TB-TC	Max abs [mm]	Mean [mm]	Stdev [mm]
max.	168.32	60.98	43.86
min.	4.38	-72.52	1.41
Mean	16.33	0.71	4.18
Stdev	7.23	5.21	1.40

TA-TC	Max abs [mm]	Mean [mm]	Stdev [mm]
max.	116.33	47.75	30.91
min.	3.45	-49.61	1.11
mean	15.71	0.13	3.86
stdev	6.93	5.69	1.40

Table 7. Statistics for the differences between the deformation time series. The statistics are computed for six pairs of teams. For each pair, the first column shows the key statistics (max., min., mean, stdev) of the maximum absolute values of the differences, the second column shows the statistics over the mean of the differences, while the third column shows the statistics for the standard deviation of the differences.

The most significant statistics from Table 7 are the means, computed over all common PS, of the means of the differences computed for each PS. These global average values, which range from -0.78 to 0.88 mm, indicate that, globally, the differences are close to zero. Another interesting global parameter is given by the mean standard deviation of the differences, which range from 3.57 to 4.34 mm. These values indicate the dispersion of the differences between time series. Other statistics, like the maximum mean difference (from 47.8 to 89.7 mm) and the maximum absolute value (from 108 to 207 mm), give an idea of the magnitude of the larger errors that can be found between the time series. These values correspond to the particular locations where at least one of the two compared teams has large errors in their estimated PS time series.

In order to get an idea of the global behaviour of the time series, a single sample T-test on the mean of the differences between pairs of time series was carried out. Using a confidence interval of 99% (α=0.01), for each pair of time series we checked the H_0 hypothesis that the mean of the differences is zero. For the six pairs, the H_0 hypothesis is accepted in a percentage that ranges from 16.9 to 19.9%. That is, in about the 80% of the cases the mean of the differences is significantly different from zero. A similar result was obtained by running a Wilcoxon signed-rank test, which represents a non-parametric alternative to the above T-test. In this case the H_0 hypothesis (median of the differences equal to zero) is accepted in a percentage that ranges from 18.01 to 20.79 %. It is worth underlining the limitation of these simple tests. Since they are based on differences between time series, they do not account for the "deformation signal", which represents the "information content" of the time series. This can be appreciated in Figures 6 and 7. The first one shows two time series that, despite their high correlation coefficient, do not satisfy the H_0 hypothesis. In this case the mean of the differences is 5.03 mm and the standard deviation is 4.22 mm. An opposite case is shown in Figure 7, which has a very low correlation coefficient, but satisfies the H_0 hypothesis. In future validation works a more in depth analysis of the time series difference should be performed, addressing key features of the time series, like the "information content" of the time series, and the PSI capability to detect specific deformation signals.

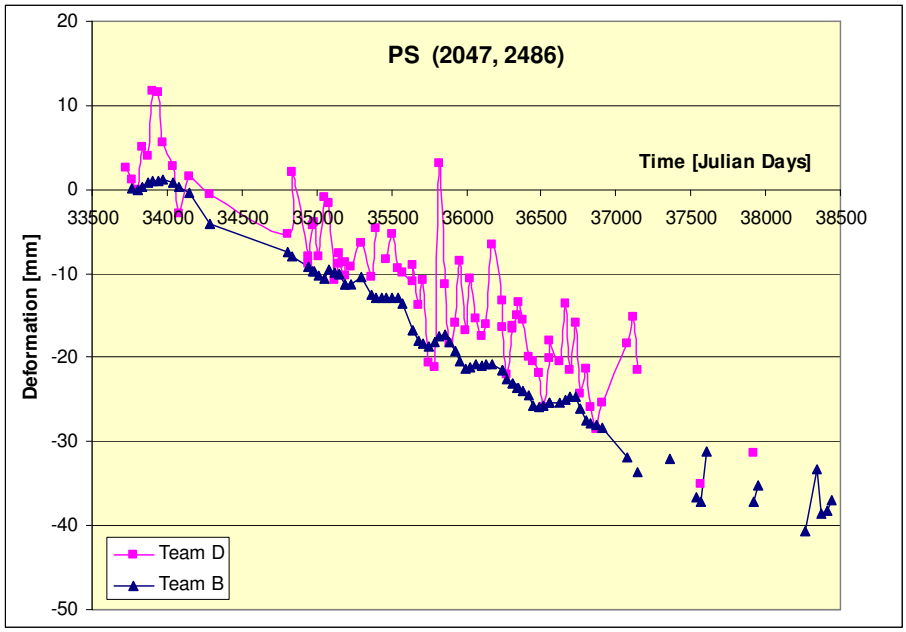


Figure 6. T-test on the mean of the differences between the time series. The mean of the differences of 5.03 mm, standard deviation of 4.22 mm, and correlation coefficient between the two time series equal to 0.91. These values were computed over 77 common samples.

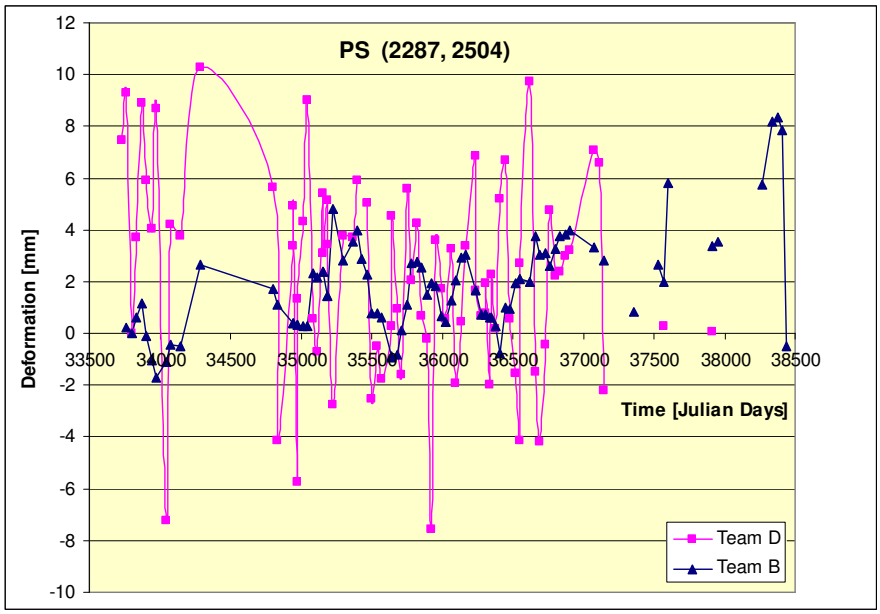


Figure 7. T-test on the mean of the differences between the time series. The mean of the differences of 0.66 mm, standard deviation of 4.16 mm, and correlation coefficient between the two time series equal to -0.06. These values were computed over 77 common samples.

We should now add some comments on the agreement between time series. Among all possible pairs of time series (the total amount is 94170), there are 13496 pairs of time series that have a correlation coefficient above 0.7, i.e. 14.3% of the total pairs. It is however worth to observe that the high correlation coefficients are mainly due to the linear deformation component of the time series. For instance, more than 99% of the time series with velocities below -2 mm/yr have a correlation coefficient above 0.7. One example of excellent agreement is shown in Figure 8. In this case the four teams estimate a cumulated deformation up to 200 mm, with rather small differences. Some exceptions are given in the last part of the series, where some problems with the Envisat data probably occur. The correlation coefficient between pairs of teams is around 0.99. Note that these values are computed by considering only the samples where both time series have an estimated deformation. Other examples of good agreement between time series are shown in Figures A.14a and b in Appendix.

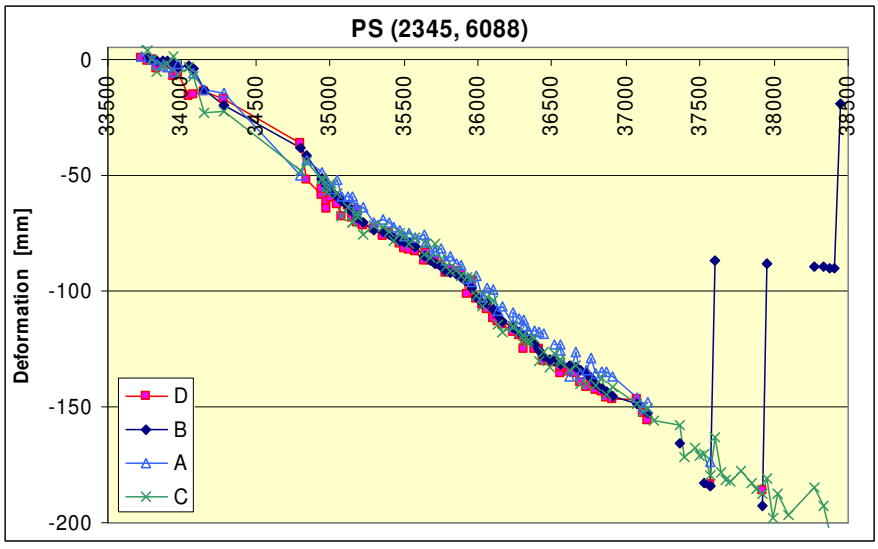


Figure 8. Example of time series where the results of the four teams show good agreement.

Taking into account that the high correlation coefficient is mainly due to the linear deformation component, it is interesting to evaluate it after removing this component. For this reason we de-trended all the time series using the deformation velocity, and computed again the correlation coefficients for all the PS pairs. 1234 pairs have a correlation coefficient above 0.7 (1.31 %), while only 4 pairs have a coefficient above 0.9 (0.004%), see two examples in Figures A.9 and A.10. These very low percentages show that, apart from the linear component, there is little in common between the time series. It is not straightforward to interpret these results. Possible explanations could be that some PS really do not have non-linear deformation, or that it is "masked" by the SAR measurement noise. In addition, one has to consider that part of the non-linear deformation of the PS can be lost during the PSI processing, and that the co-registration errors between the PS datasets surely contribute to lower the correlation coefficients.

We now consider the cases where the time series show important discrepancies. In some cases the discrepancies reflect the big differences in the deformation velocity estimated by the teams. For instance, as already mentioned in Section 3 there are between 0.07% (TB – TA) and 0.25% (TD – TC) of PS with absolute differences above 5 mm/yr. There is a subset of time series that shows a particular behaviour: they are negatively correlated. This means that when one team estimates a subsidence, in the same period the other team estimates an upsidence. This behaviour is probably due to some bug in the PSI processing. It is worth noting that only a rather small set of time series pairs show the above behaviours. For instance, there are 100 pairs that have a correlation coefficient below -0.7. This represents about the 0.1 % of total. Finally it is worth mentioning that some time series include outliers in the form of deformation peaks, which are usually associated with a given temporal sample. Even though we did not perform an ad hoc outlier detection analysis on the time series, we can say that this type of error is not rare. This suggests an **input for future PSI validation experiments**:

Suggest that the teams **properly filter the outliers** in the time series**.**

Below other inputs for future PSI validation experiments are discussed.

1) In order to compare two time series, a common temporal reference (origin) is needed. In this work the common reference was fixed in the 3rd temporal sample, simply because it was the first one where the four teams have data (the teams did not use exactly the same set of SAR images). The disadvantage of this procedure is that a noisy measurement in the third sample affects the entire time series. An example is shown in Figure A.11 of the Appendix: the TD time series is biased because the origin of the time series coincides with a noisy sample. This provides an **input for future PSI validation experiments**:

Fix the rules for appropriately **choosing the origin of the time series** a priori.

2) Though the coherences provided by the teams were not considered in this study, we will briefly mention that there are remarkable differences in the way the temporal coherence of the PS is computed. An example is shown in Figure A.12 in the Appendix: over the same PS the four teams estimate rather different coherence values, which range from 0.66 to 0.96. These differences represent a limitation for the inter-comparison of PSI results that come from different teams. This can suggest a further **input for future PSI validation experiments**:

Homogenize the way of computing the temporal coherence, in order to get a **quality index** that can be used to inter-compare the PSI results of different teams.

5. CONCLUSIONS

This document describes the results of the Provence inter-comparison, which are based on the PS datasets generated in the PSIC4 project. The PSIC4 test site is located around the mining area of Gardanne, in Provence (southern France). The inter-comparison concerns four of the height PSI teams involved in the PSIC4 project, which are among the five OSPs of the Terrafirma project: Altamira Information, DLR, Gamma Remote Sensing and TRE. Since this study has been exclusively based on inter-comparison, it is important to underline that its results should not be confused with those coming from validation experiments, where the PS products are compared against ground truths. In the inter-comparison the coherences provided by the teams were not considered.

The Provence inter-comparison has specific characteristics which make it different from the analysis performed in PSIC4. Firstly, it is only based on the original PS data, defined in the radar image space, while PSIC4 concerned both geocoded PS products and data in the original radar coordinates. Then, it concerns a remarkably larger area with respect to PSIC4, 925 vs. 100 km^2. It is only focused on inter-comparison, while PSIC4 involved both validation and inter-comparison activities. Furthermore, it concerns only deformation velocities and deformation time series, while in the PSIC4 project other aspects of the PSI products were analysed. Finally, the Provence inter-comparison is restricted to four of the five OSPs of Terrafirma, while PSIC4 involved the analysis of the results of eight PSI teams.

These conclusions include three parts. Firstly, the key inter-comparison results are summarized. Then the main inputs for new validation experiments are listed. Finally, in the last part these results are compared with those of the PSIC4 project.

The most relevant results related to the deformation velocity maps are summarized below.

1) The statistics of the velocity differences provide information on the global inter-comparison behaviour of the teams. One of the most important results concerns the standard deviations of the velocity differences, which range from 0.58 to 0.80 mm/yr (Table 2). These values are confirmed by the analysis of the distribution of the differences: in the statistics that include Gardanne (Table 4), at least 77% of the common PS have absolute velocity differences below 1 mm/yr.
2) The means of the velocity differences range from -0.37 to 0.18 mm/yr (Table 2). Even though these are rather small values, a T-test shows that five of six team pairs have means of the differences significantly different from zero, with a confidence interval of 99%. This can be due to small error when referring the PS datasets to the same stable area, or other sub-millimetric residual systematic errors in the velocities maps.
3) All the inter-comparison statistics computed on the global test area, including the Gardanne mining area, are similar to those computed by excluding the mining area, where the PS validation gives its worst results. That is, over the mining area of Gardanne, the quality of the PS inter-comparison does not strongly decrease. A partial explanation of this is that over the most difficult areas, i.e. those with big deformations, there is a low PS density.
4) The statistics on the velocity differences show that there are PS with large absolute velocity differences (see maximum and minimum values in Table 2). However, it is worth underlining that these PS represent a rather small fraction of the entire PS dataset.
5) The statistics for the velocity differences, which were computed for different classes of velocity, show that the best PS differences correspond to the velocity around zero, while the dispersion of the differences increases when the module of the velocity increases.

The most relevant results related to the deformation time series are summarized below.

1) The global behaviour of the time series is summarized by the average, computed over all common PS, of the means of the differences computed for each PS (Table 7). These average values, which range from -0.78 to 0.88 mm, indicate that globally the differences are close to zero.
2) Another important global parameter is given by the mean standard deviation of the differences, which ranges from 3.57 to 4.34 mm. These values indicate the dispersion of the differences between time series.
3) In order to get an idea of the global behaviour of the time series, a single sample T-test on the mean of the differences between pairs of time series has been carried out. With confidence interval of 99%, the H_0 hypothesis that the mean of the differences is zero is accepted in a percentage that ranges from 16.9 to 19.9%. That is, in about the 80% of the cases the mean of the differences is significantly different from zero. A similar result was obtained by running a Wilcoxon signed-rank test, which represents a non-parametric alternative to the above T-test. However, in this study the limitations of these tests have been discussed, suggesting that in future validation works a more in depth analysis of the time series differences should be performed. The analysis should address key features, like the "information content" of the time series, and the PSI capability to detect specific deformation signals.
4) 14.3% of time series pairs have a correlation coefficient above 0.7. However it is worth to observe that the high correlation coefficients are mainly due to the linear deformation component of the time series. For instance, more than 99% of the time series with velocities below -2 mm/yr have a correlation coefficient above 0.7. For this reason we analysed the correlation coefficient between time series after removing the linear deformation component. 1.31% of the pairs have a correlation coefficient above 0.7. This low percentage shows that, apart from the linear component, there is little in common between the time series. Possible explanations could be that some PS really do not have non-linear deformation, or that it is "masked" by the SAR measurement noise. In addition, one has to consider that a part of the non-linear deformation of the PS can be lost during the PSI processing, and that the co-registration errors between the PS datasets surely contribute to lower the correlation coefficients.
5) The final result concerns the time series that show important discrepancies. A small fraction of them have big differences in the deformation velocity. For instance, there are between 0.07% and 0.25% of PS with absolute differences above 5 mm/yr. Furthermore, there is a small subset of time series that are negatively correlated. This behaviour is probably due to some bug in the PSI processing. Finally, some time series include outliers in the form of deformation peaks, which probably could be avoided adopting an appropriate data filtering.

Some of the key findings of this inter-comparison work can represent a valuable source of inputs for carrying out new PSI validation experiments. The most relevant inputs are briefly listed below.

1) A key pre-processing step needed for the PS inter-comparison has been the co-registration of the PS datasets. This co-registration suffers an important limitation, because it has been realized a posteriori, using only the PS tables. Much better co-registration performances could have been achieved by using the SLC images of the teams. For future projects we suggest planning the experiment accurately in order to get a very accurate dataset co-registration in the radar space, directly from the PSI teams.

2) Another pre-processing step has been to refer the PS datasets to the same area. For future validation projects we suggest defining the stable reference area a priori and requiring the PSI teams to refer their results to this area.

3) The time series sometimes include outliers in the form of deformation peaks. These outliers degrade the quality of the deformation time series. We therefore believe that for future work the validation should be performed on properly filtered time series, avoiding the above outliers.

4) In order to inter-compare the time series, a common temporal reference has to be fixed. Since this operation can bias the inter-comparison, the way of fixing the origin of the time series should be defined before running the PSI processing.

5) There is a remarkable difference in the way the temporal coherence of the PS is computed. This represents a limitation for the inter-comparison of PSI results. If possible, the way of computing the temporal coherence should be homogenized, in order to get a quality index that can be used to inter-compare the PSI results of different teams.

Finally, we briefly compare below the results of the inter-comparison with those of the PSIC4 project. As mentioned above, there are key differences between the PSIC4 project and the Provence inter-comparison. However, the results of the two projects are not contradictory. In some cases they simply show different complementary aspects. For instance, the Provence inter-comparison is largely based on data (outside the Gardanne mining area), which were simply not analysed in PSIC4. In other cases the results are rather similar, as it is briefly discussed in the following three examples.

1. The first example concerns the velocity shifts to fix the reference area in both projects. A part from a global shift of about 0.5 mm/yr, due to the fact that different reference areas were used, the relative differences between the teams, e.g. $mean_{TA}$ - $mean_{TB}$, are very close to the values assessed in PSIC4. In other words, the values estimated in this study are basically consistent with those of PSIC4.

2. The second example concerns the global statistics of the velocity inter-comparison, see Table A.1. The mean differences are similar for the pairs that do not include TC. For the other pairs there is a shift of about 0.3-0.4 mm/yr, which however is consistent with other results shown in this document, e.g. Table 4 and 6. The standard deviations computed in the Provence inter-comparison are smaller than those of PSIC4 because the latter include only the statistics computed over the mining area of Gardanne, where the PS validation gives its worst results.

3. The last example concerns the time series, see Figure A.13. The plot coming from the PSIC4 project (above) shows one levelling and four PS time series. In the Provence inter-comparison (below), three of these time series show the same behaviour. This example confirms that the Provence inter-comparison and the PSIC4 project show consistent results.

ANNEX 1

A1.1 Reference of the 4 datasets to the same stable area

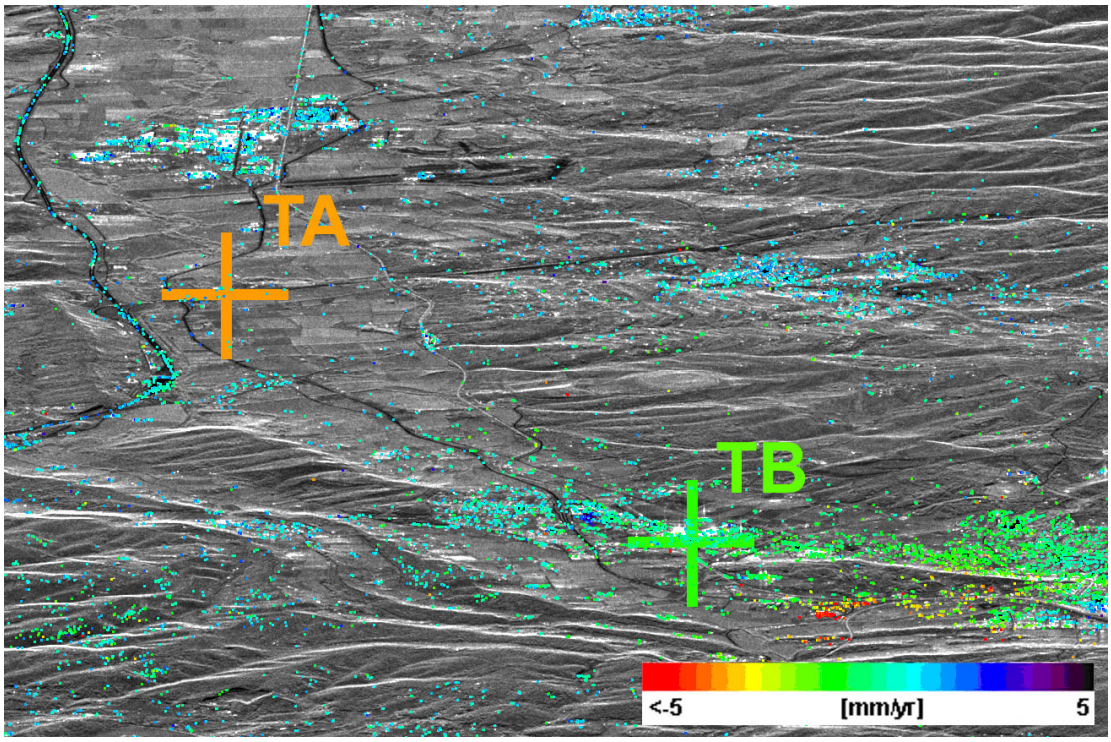


Figure A.1 Velocity map corresponding to the team A. The zoom includes the two reference points corresponding to the team B and A. Covered period: 92-00

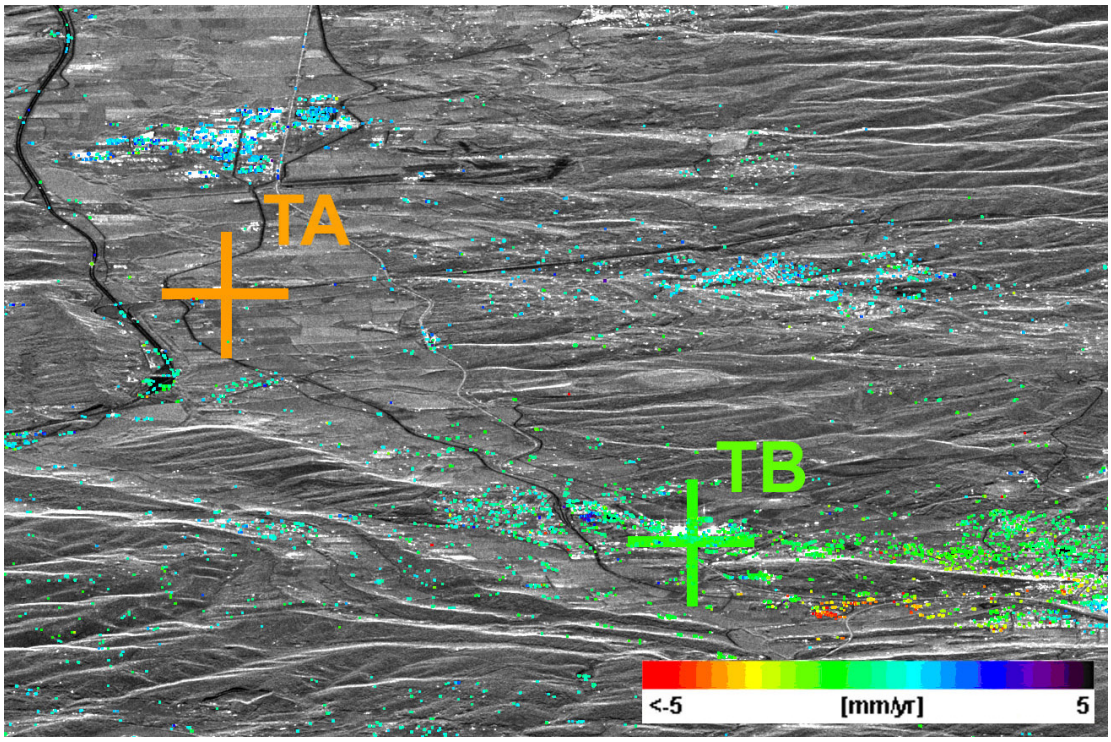


Figure A.2 Velocity map corresponding to the team B. The zoom includes the two reference points corresponding to the team B and A. Covered period: 92-00

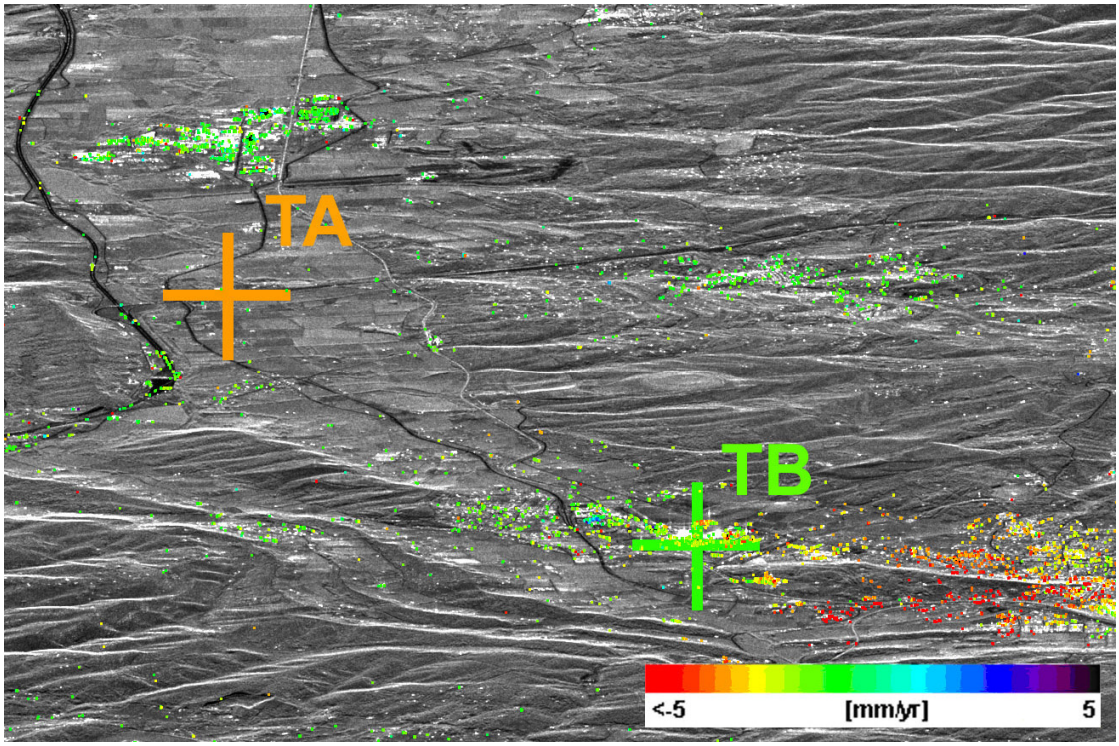


Figure A.3 Velocity map corresponding to the team C. The zoom includes the two reference points corresponding to the team B and A. Covered period: 92-00

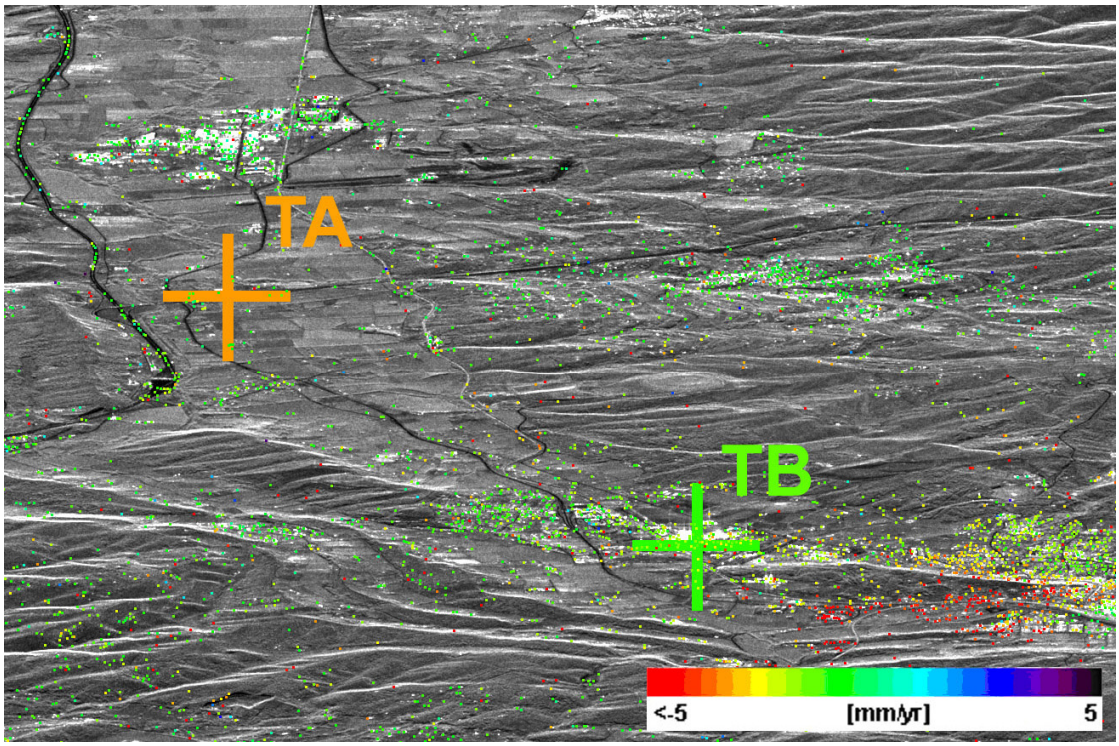


Figure A.4 Velocity map corresponding to the team D. The zoom includes the two reference points corresponding to the team B and A. Covered period: 92-00

A1.2 Velocity maps

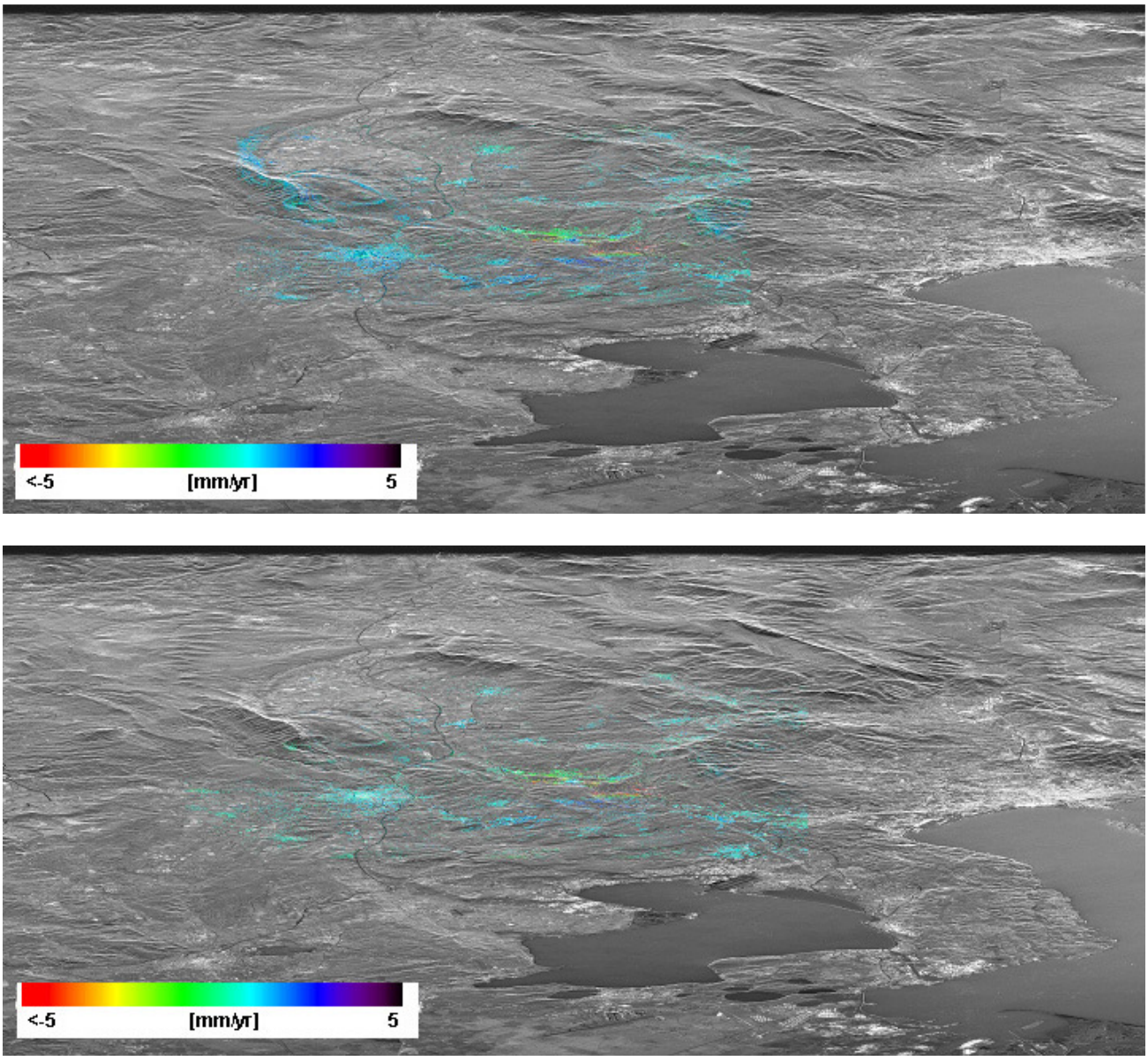


Figure A.5 Original velocity maps, before referring them to the same reference area, corresponding to the teams A (above) and B (below). Covered period: 92-00

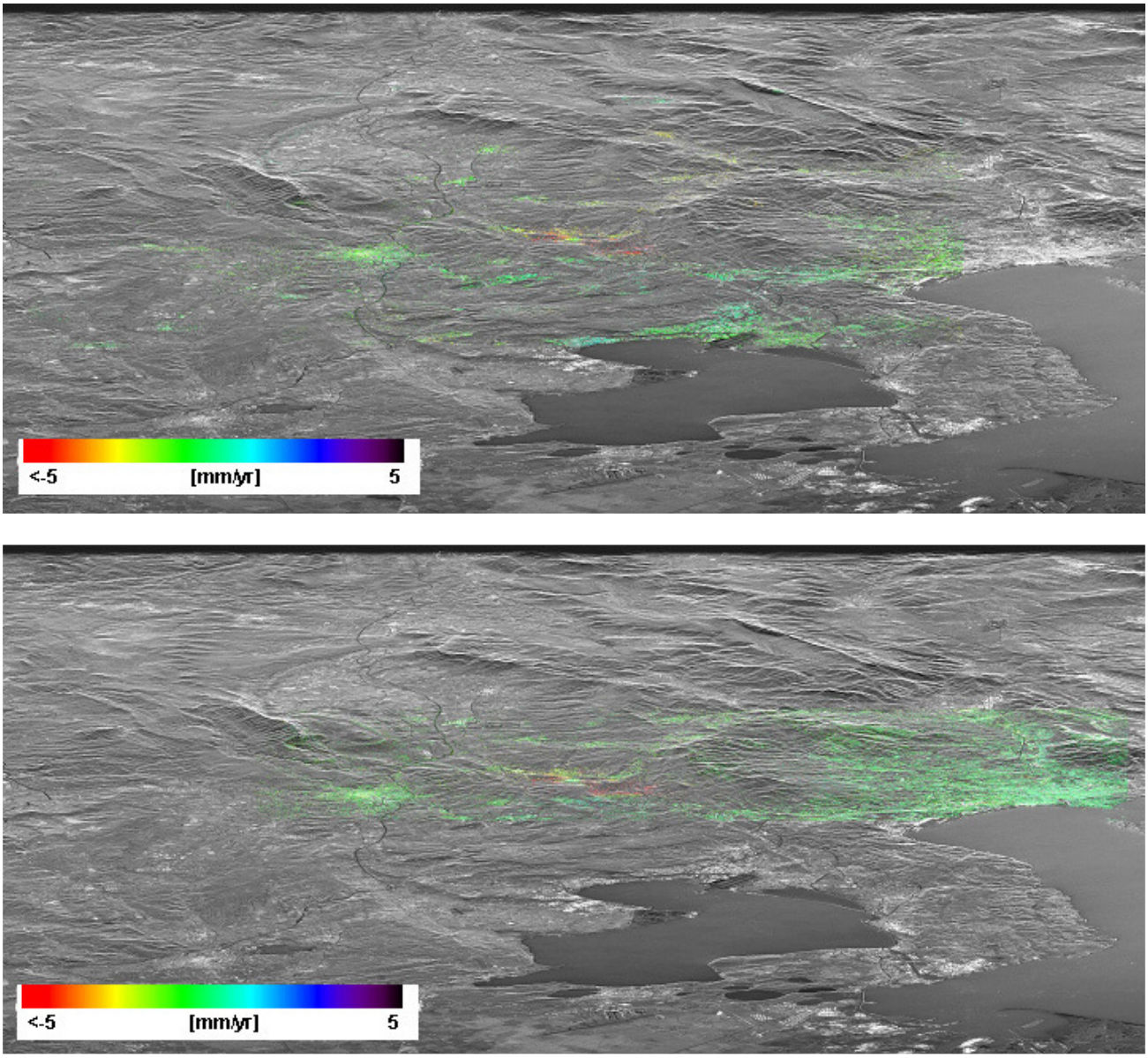


Figure A.6 Original velocity maps, before referring them to the same reference area, corresponding to the teams C (above) and D (below). Covered period: 92-00

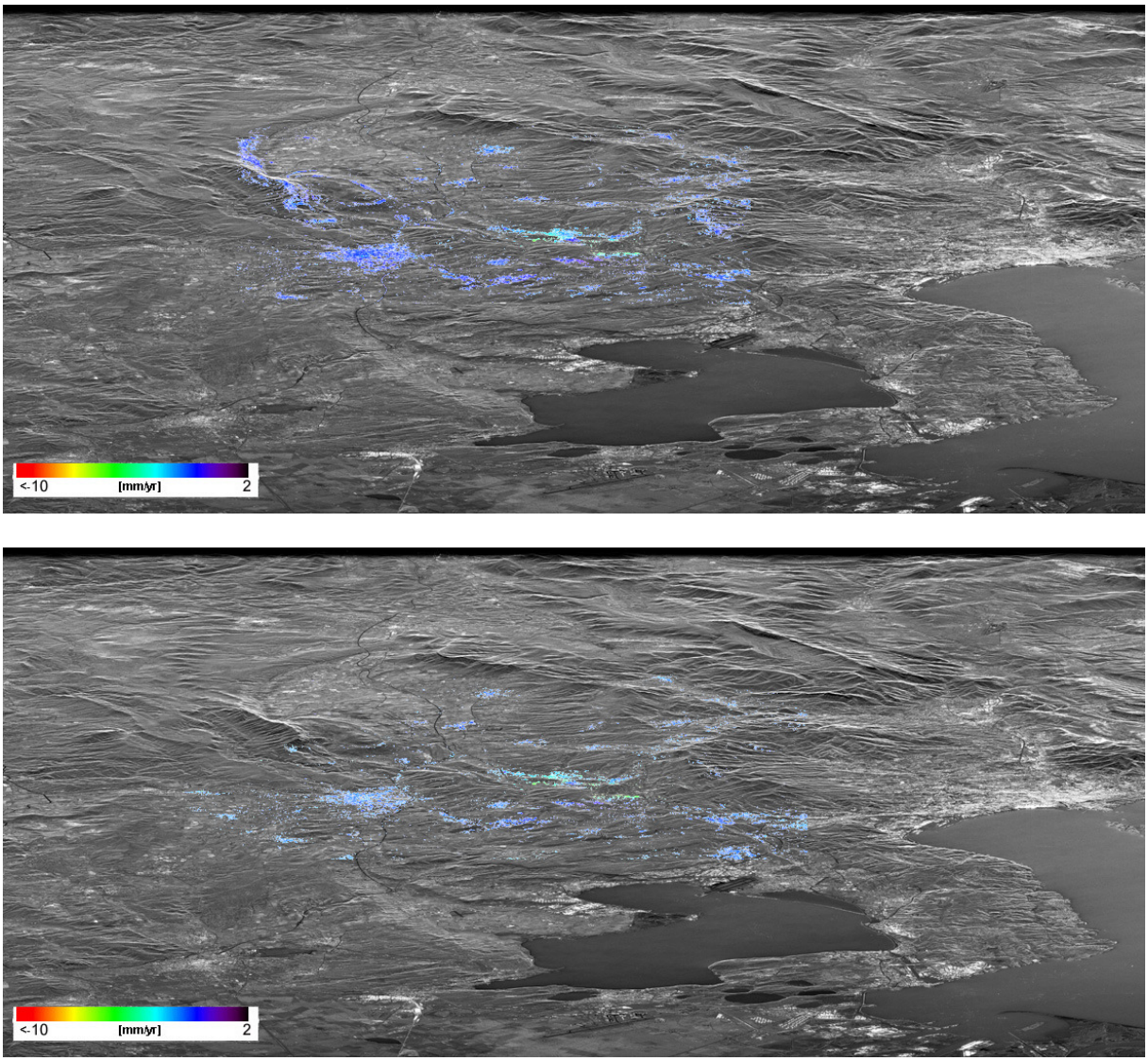


Figure A.5 Original velocity maps, after referring them to the same reference area, corresponding to the teams A (above) and B (below). Covered period: 92-00

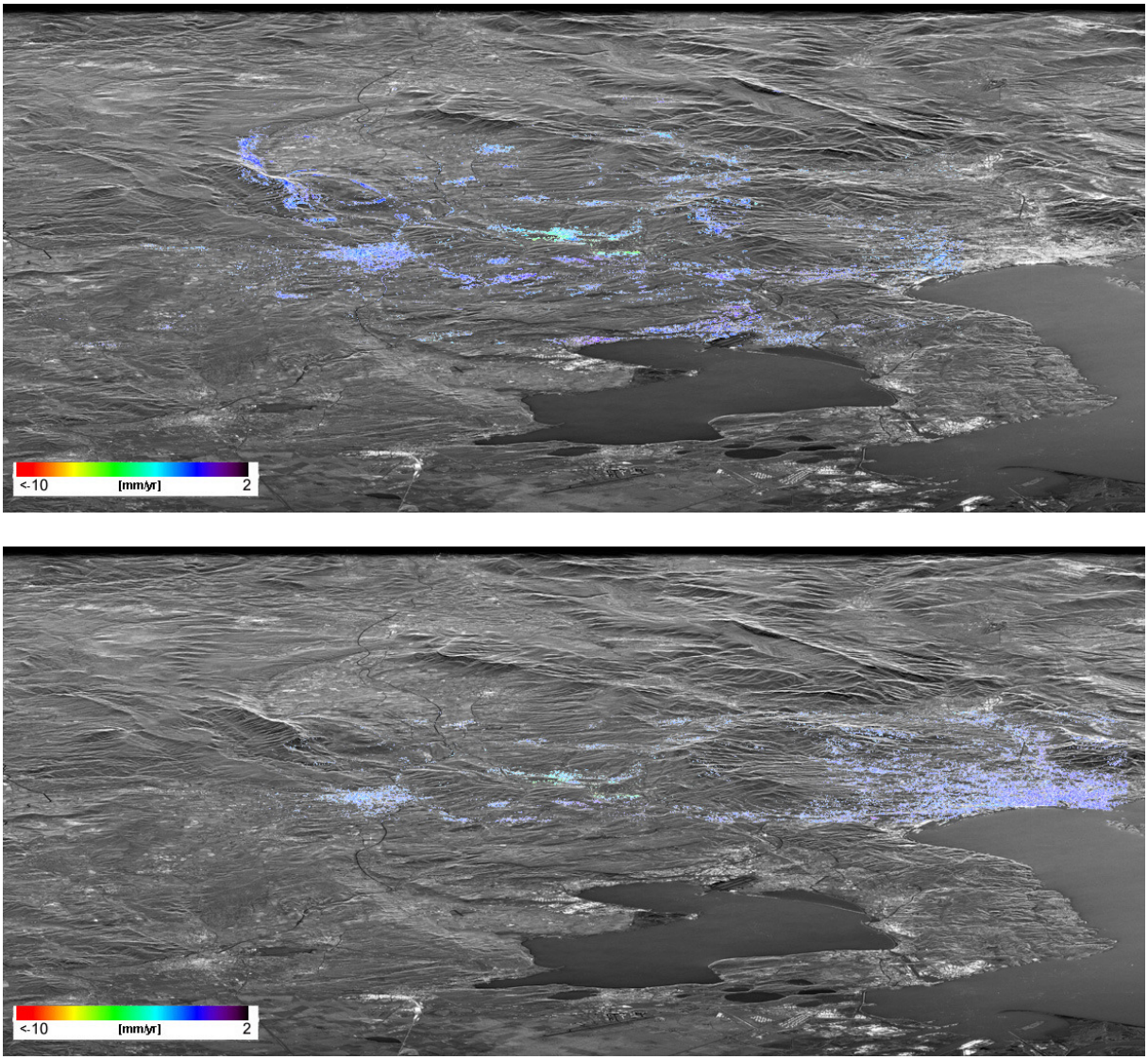


Figure A.6 Original velocity maps, after referring them to the same reference area, corresponding to the teams C (above) and D (below). Covered period: 92-00

A1.3 Traffic light deformation maps

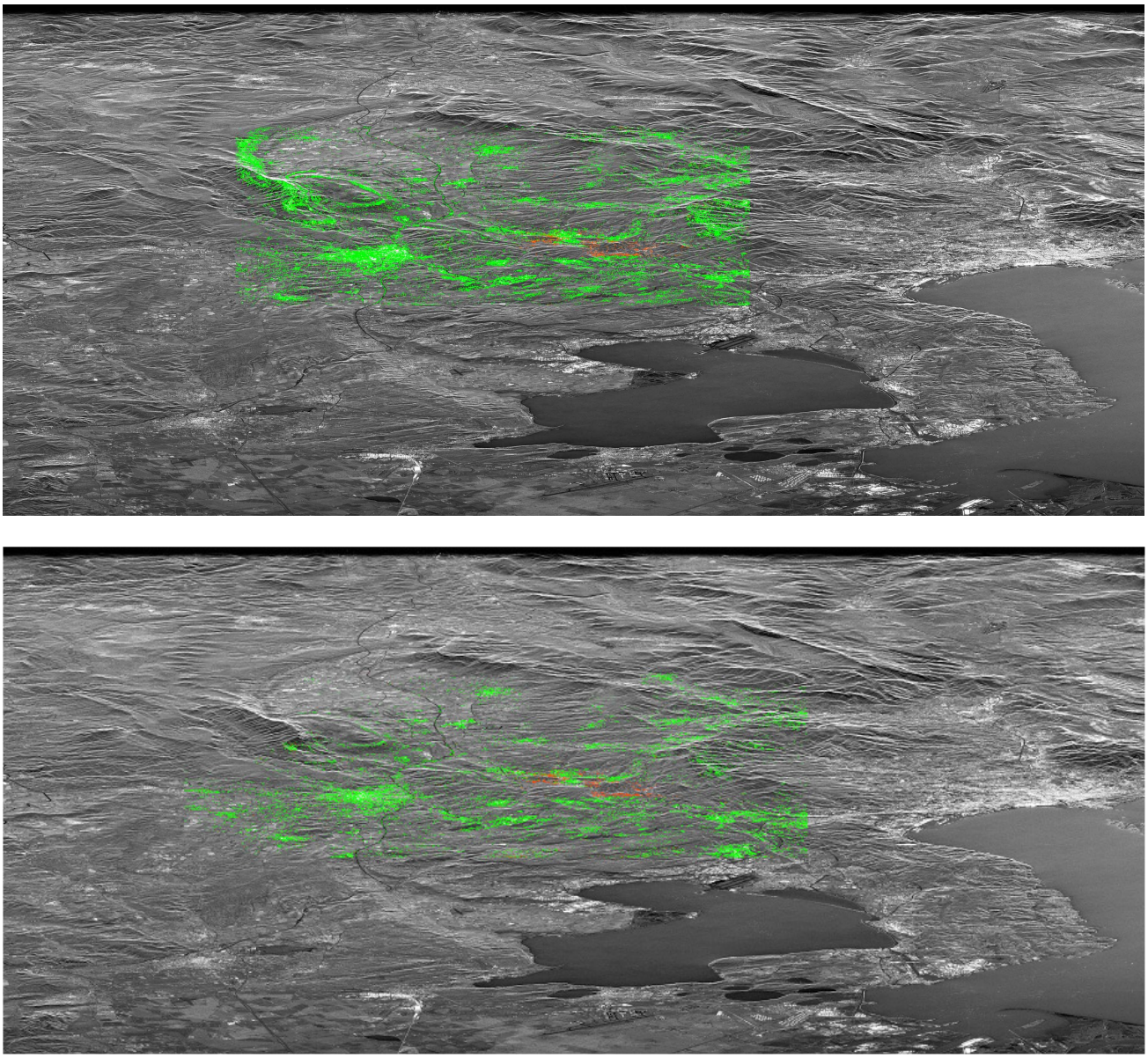


Figure A.7. Traffic light maps corresponding to teams A (above) and B (below).

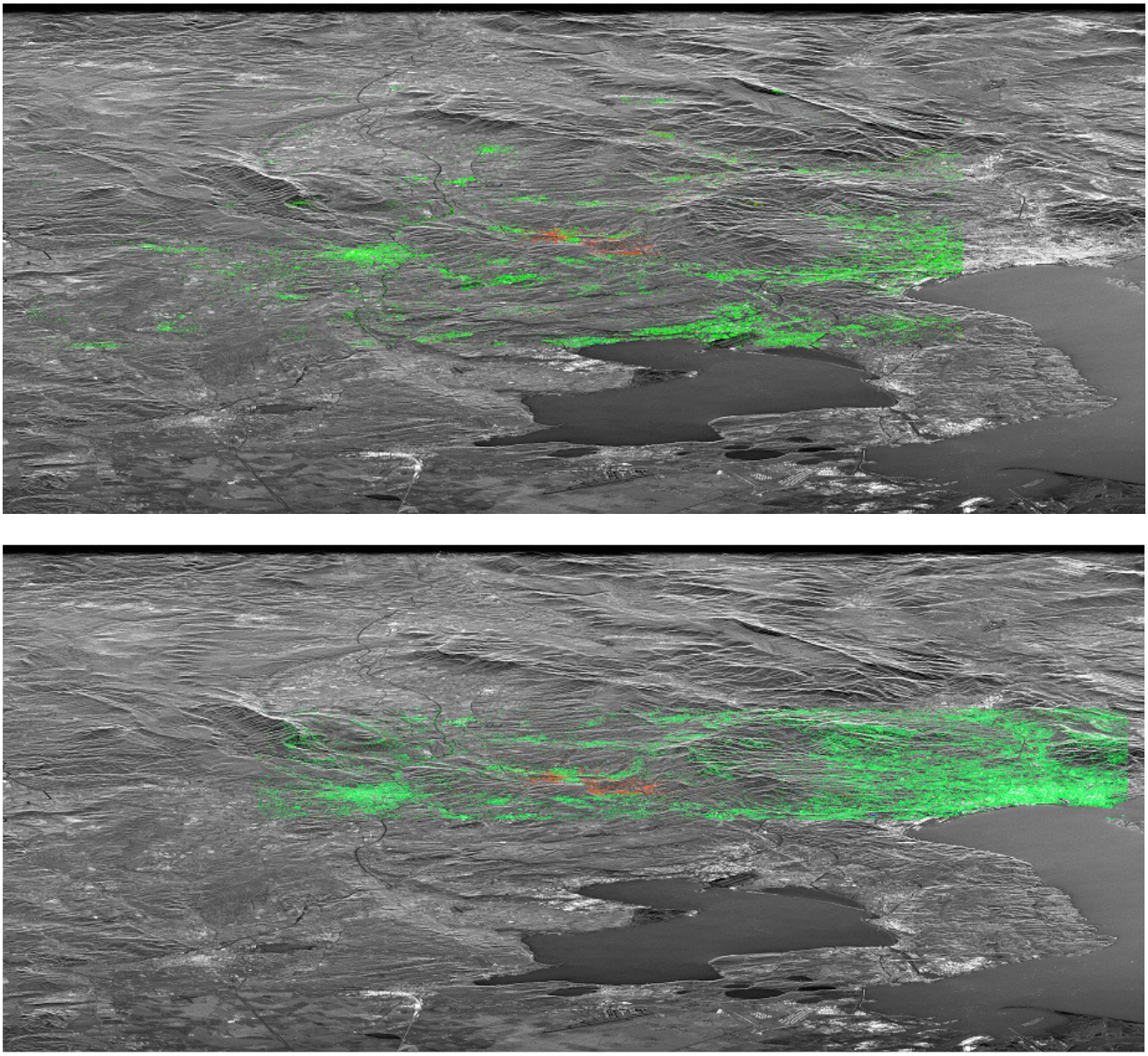
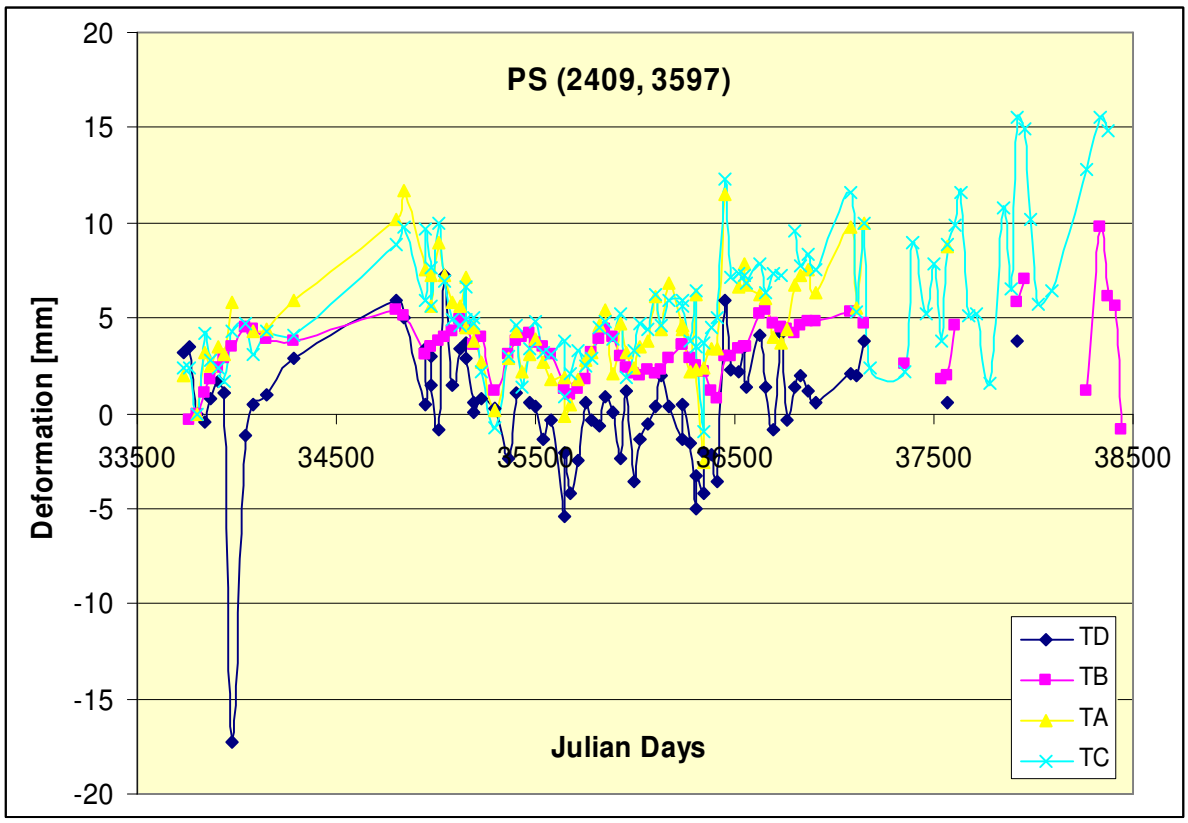


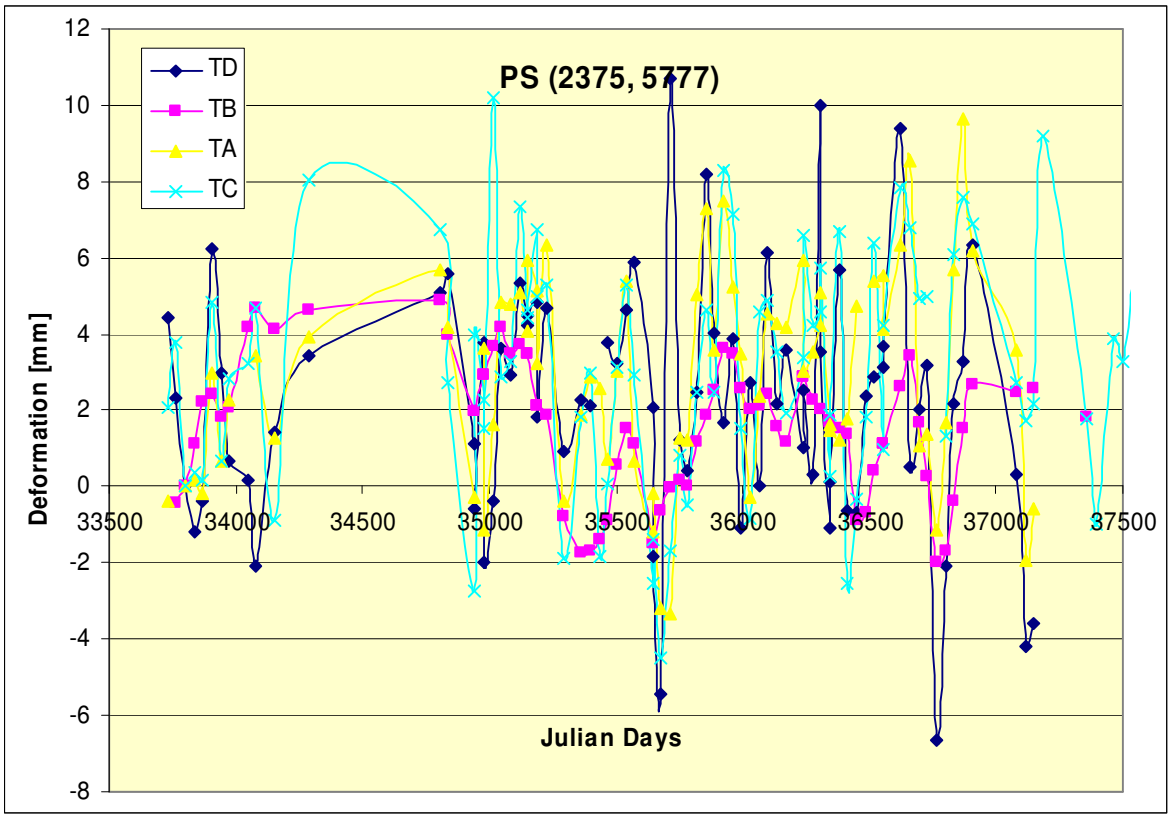
Figure A.8. Traffic light maps corresponding to teams C (above) and D (below).

A1.4 Analysis of de-trended time series



Pair.	Corr. Coeff.
D-B	0,385
D-A	0,457
D-C	0,412
B-A	0,582
B-C	0,631
A-C	0,912

Figure A.9. Example of de-trended time series. One pair shows a high correlation coefficient. There are only four pairs, out of 94170, that have a correlation coefficient above 0.9 (0.004%).



Pairs	Corr. Coeff.
D-B	0,222
D-A	0,445
D-C	0,393
B-A	0,483
B-C	0,852
A-C	0,673

Figure A.10. Example of de-trended time series. One pair shows a relatively high correlation coefficient.

A1.5 Inputs for future validation experiments

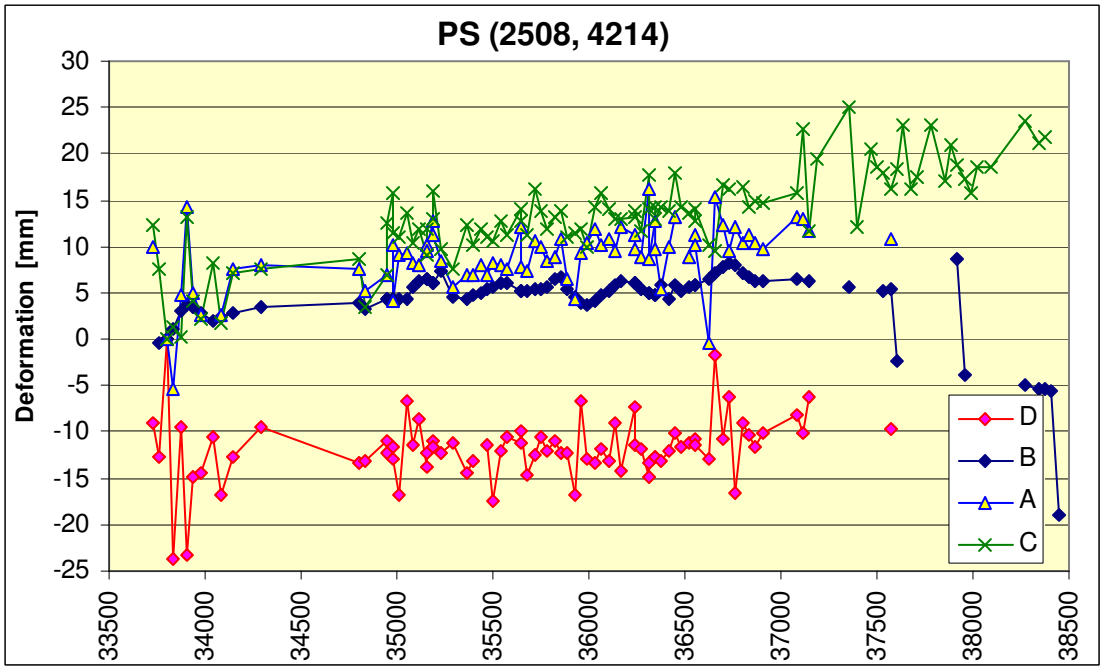


Figure A.11. The TD time series is shifted (say, biased) because the origin of the time series coincides with a noisy sample. This shows the importance of appropriately choose the temporal origin for all the time series.

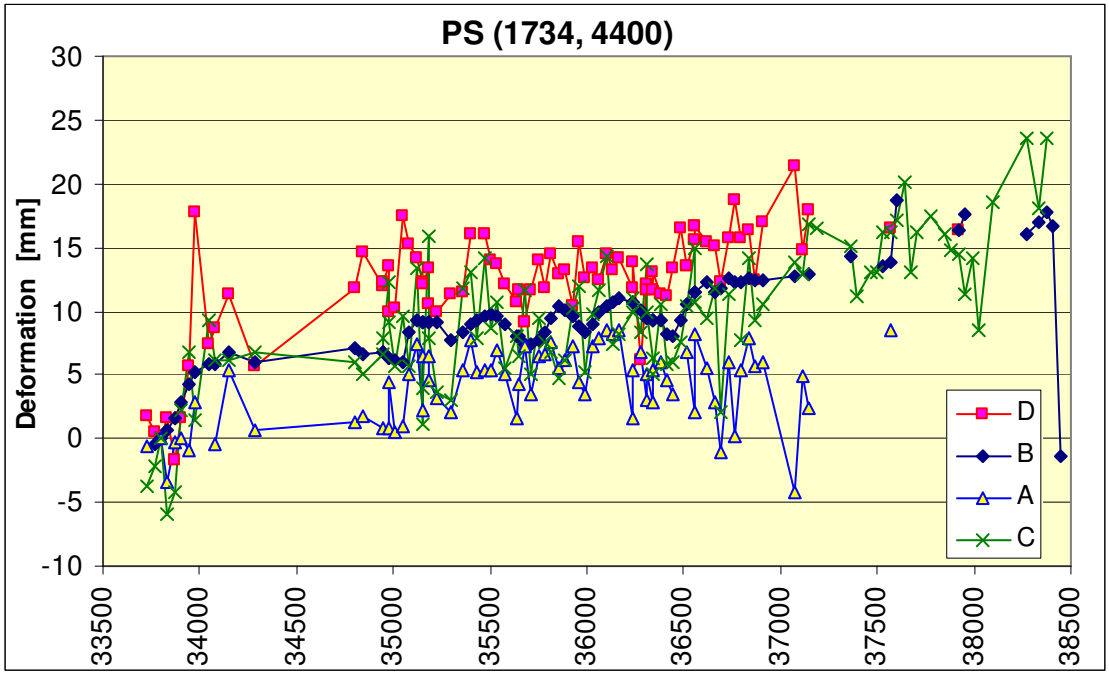
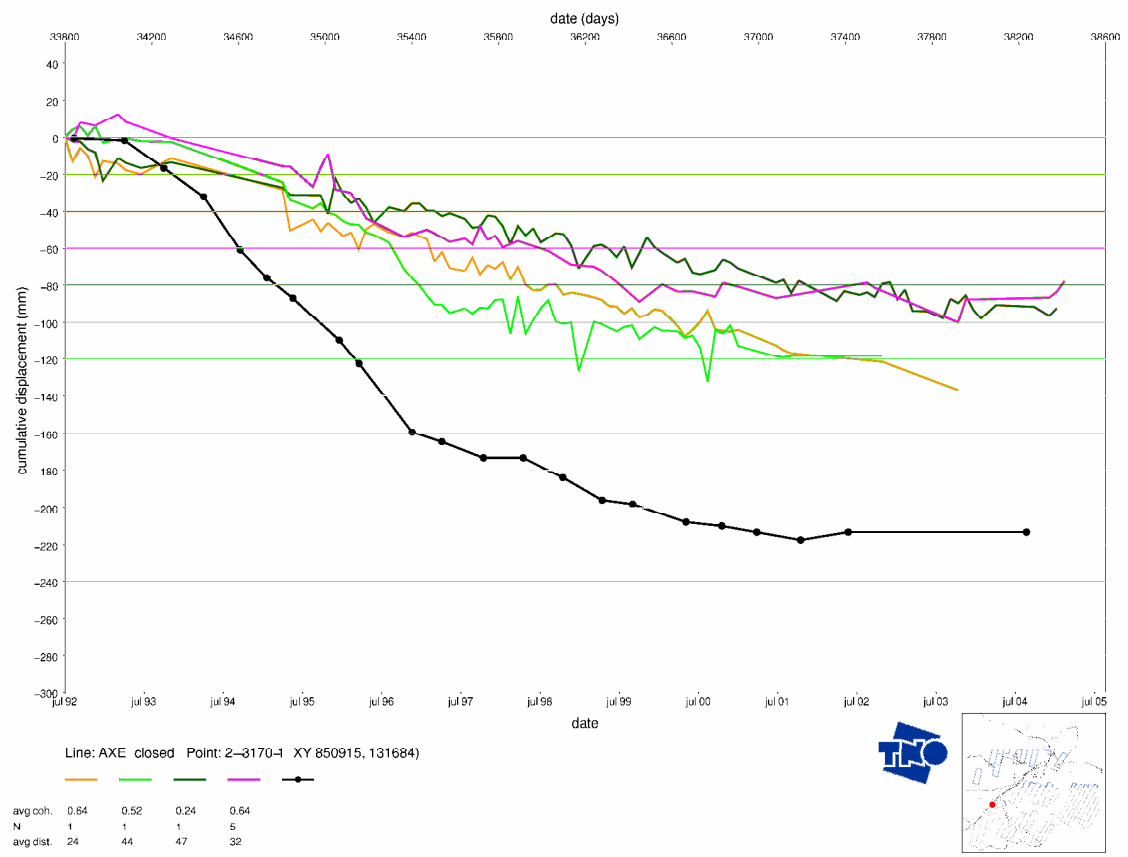


Figure A.12. Four time series that exhibit different temporal coherence: TA = 0.84, TB = 0.96, TC =0.66, and TD = 0.79.

A1.6 Result comparison with PSIC4

	Provence inter-comparison		PSIC4 project	
[mm/yr]	**Mean**	**Stdev**	**Mean**	**Stdev**
TD-TB	0.18	0.70	0.12	0.87
TD-TA	-0.09	0.63	-0.13	1.06
TD-TC	-0.02	0.77	0.27	1.19
TB-TA	-0.32	0.58	-0.24	0.63
TB-TC	-0.37	0.80	0.20	0.87
TA-TC	-0.08	0.79	0.44	1.01

Table A.1. Comparison of the global statistics of the Provence inter-comparison and the PSIC4 project. The statistics of the Provence inter-comparison are the same shown in Table 2, while those of PSIC4 come from the task named "Inter-comparison of velocity maps". The mean differences of the three pairs that do not include TC are similar, while those that include TC show a shift of about 0.3-0.4 mm/yr, which is consistent with other results shown in this document, e.g. see Table 4 and 6. The standard deviations computed in the Provence inter-comparison are smaller than those of PSIC4 because the latter ones only include the statistics computed over the mining area of Gardanne, where the PS validation gives its worst results.



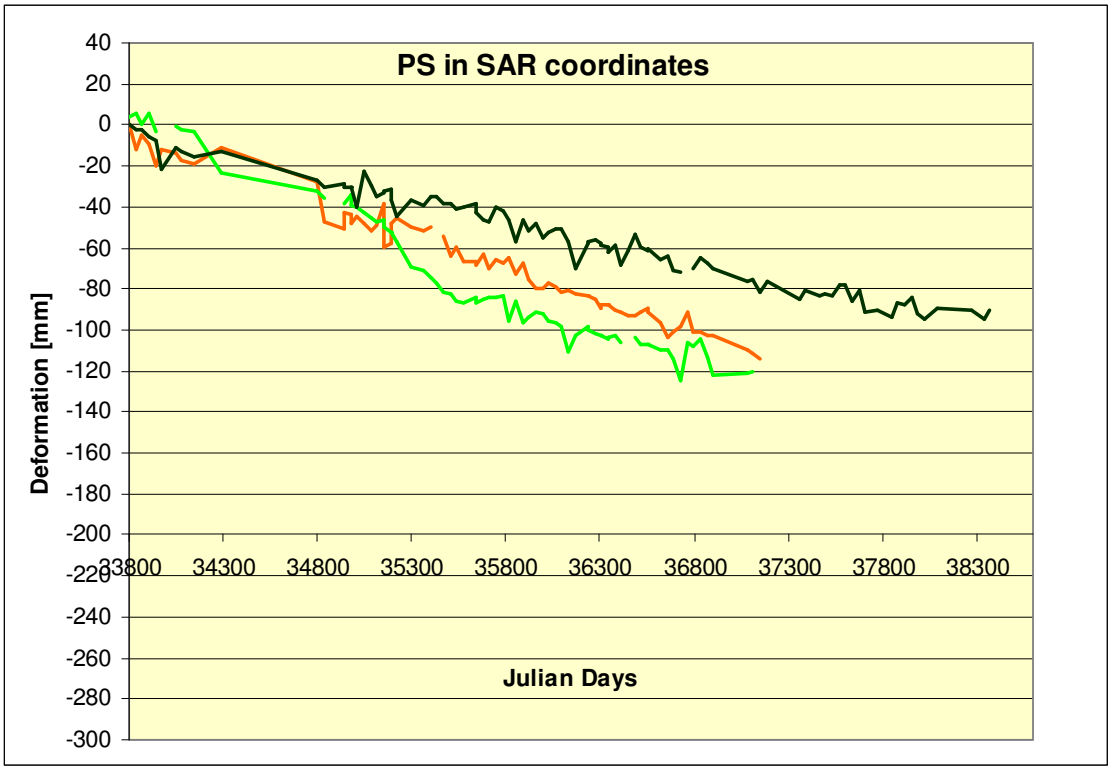


Figure A.13. Plot showing one levelling and four PS time series coming from the PSIC4 project (above). Three of the above time series, computed in the radar coordinates in the Provence inter-comparison (below). The time series show the same behaviour.

A1.6 Examples of good agreement

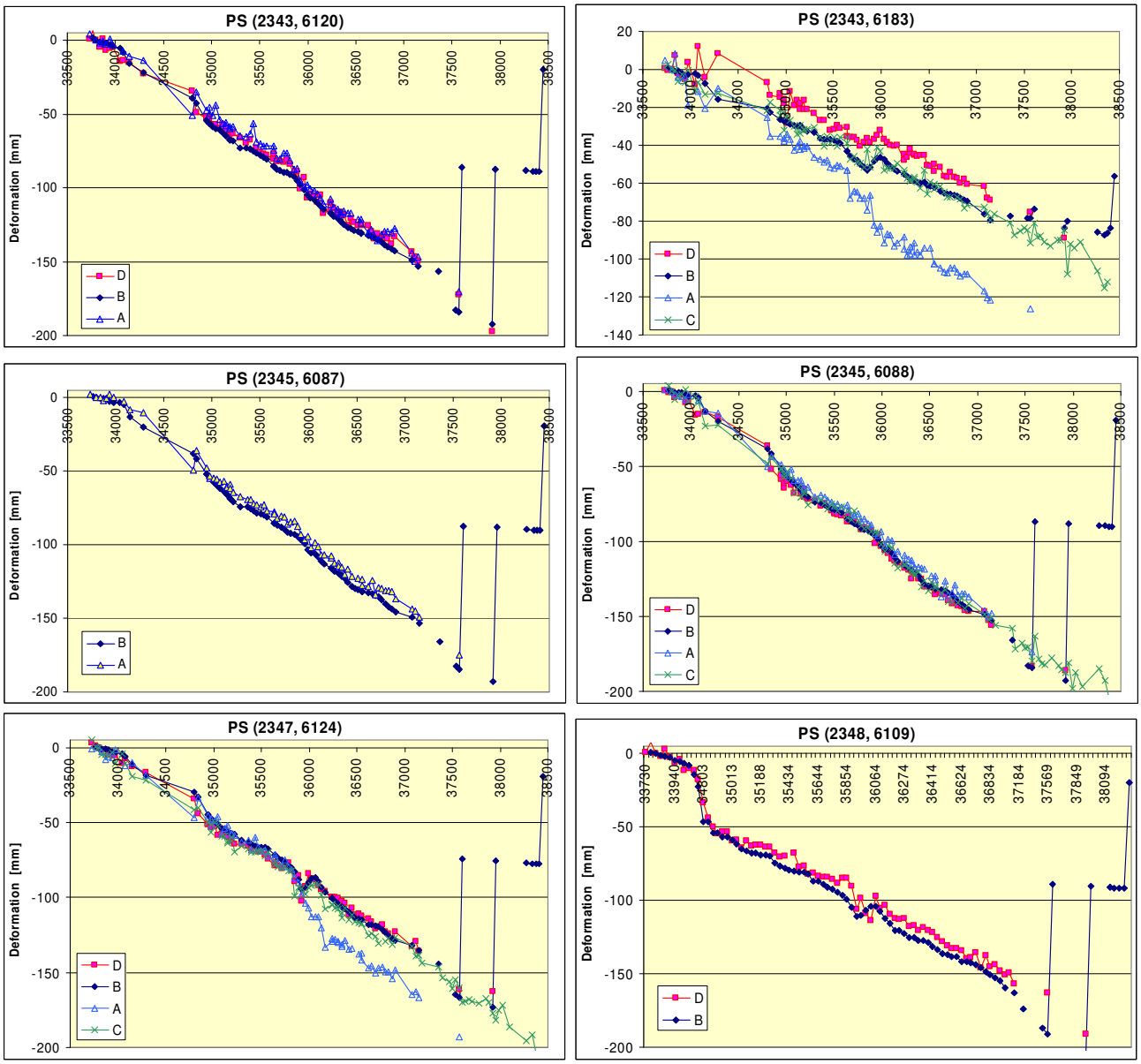


Figure A.14a. Example of time series that show a good agreement.

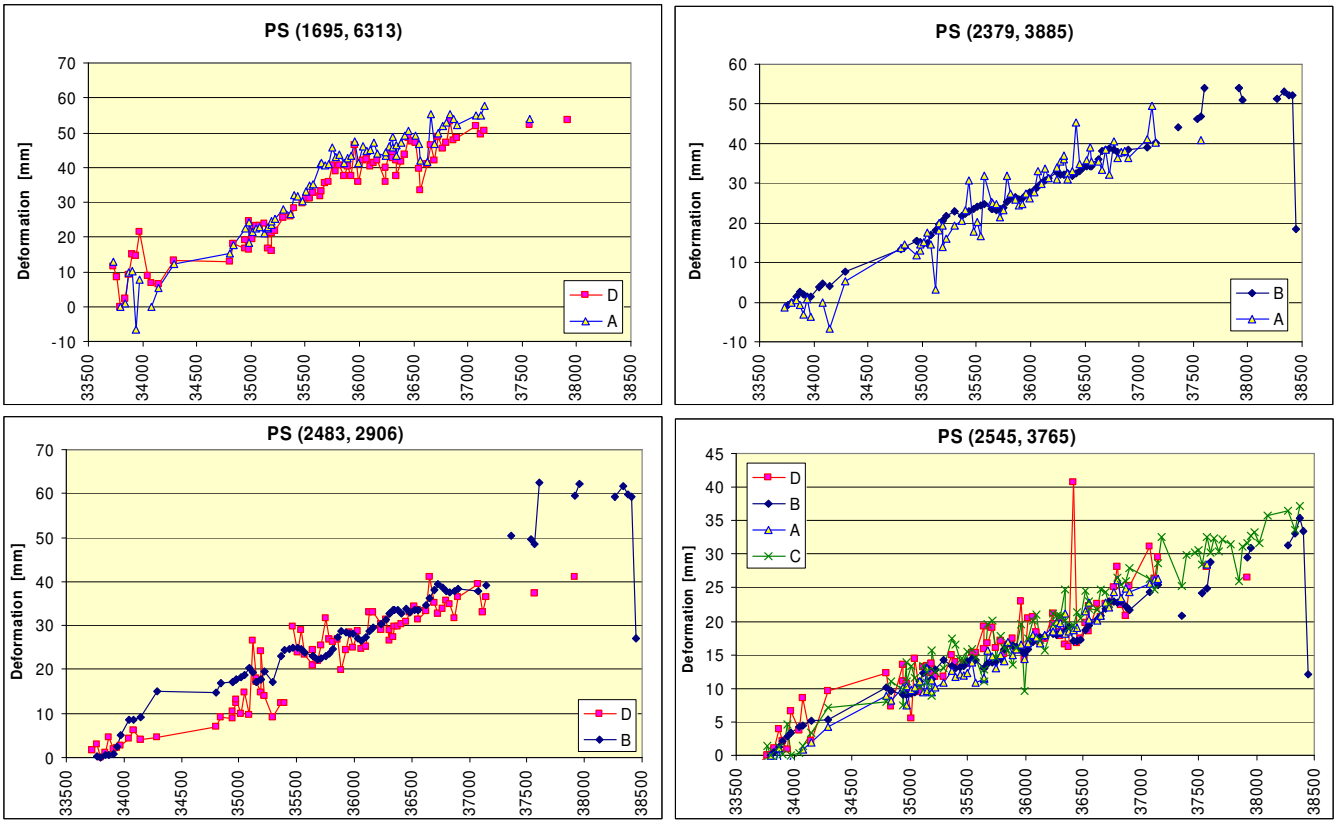


Figure A.14b. Example of time series that show a good agreement.