GMES TERRAFIRMA: VALIDATION OF PSI FOR USERS
RESULTS OF THE PROVENCE INTER-COMPARISON


(1) Institute of Geomatics, Parc Mediterrani de la Tecnologia, Av. del Canal Olímpic, s/n E-08860,Castelldefels (Spain), michele.crosetto@ideg.es, marta.agudo@ideg.es

(2) Nigel Press Associates Ltd, Crockham Park, TN8 6SR, Edenbridge (United Kingdom), ren@npagroup.com

(3) British Geological Survey, Nicker Hill, Keyworth, Nottingham, NG12 5GG (United Kingdom), shm@bgs.ac.uk

ABSTRACT

The GMES Terrafirma project provides a ground movement hazard information service for Europe, which is based on the exploitation of time series of SAR data and the Persistent Scatterers Interferometry (PSI) technique. Since PSI is a relatively new technique, which can potentially be used in a wide range of applications, the validation of its land deformation measurement products plays a key role to increase the acceptability of the technique and to establish a long-term market. In the Terrafirma project a special interest is given to the validation issues of PSI deformation measurement, and, in particular, to the validation of the Level 1 products, which are basic products providing raw PSI ground motion measurements. This paper describes the main results of the inter-comparison of the products of four Terrafirma PSI teams involved in the PSIC4 project. The paper describes the pre-processing tasks performed to achieve homogeneous and co-registered datasets, the analysis of the deformation velocity maps, and the study of the deformation time series.

1. INTRODUCTION

Within the PSIC4 (Persistent Scatterers Interferometry Codes Cross Comparison and Certification for Long term differential interferometry) project, eight teams specialized in PSI processed a large stack of SAR images over the same test site [1]. The site was located around the mining area of Gardanne, in Provence (France). Though the areas covered by the different PSI teams are considerably larger than the Gardanne mining area, see Fig. 1, the PSIC4 project was mainly focused on this area, where most of the ground deformations occur and the ground truths are available. The analysis 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. An additional motivation to perform this study was to collect inputs for the preparation of a new validation experiment, to be carried out as part of the Terrafirma project. The focus of this study is on the global inter-comparison performances of four TF OSPs. The inter-comparison studies how consistent are the results between different teams? It is worth emphasizing that all the results discussed in this paper 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”, avoiding 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 km². This is the area where the bigger land deformations occur. 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 km², see Fig. 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.

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 OSP, 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.

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, two pre-processing steps were performed. The first one, homogenizing the datasets, 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. The second one, was referring the four PS datasets to the same stable area.

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 Fig. 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. However, for at least one team there were two shifts, separated by a difference of one pixel, which 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.

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 Fig. 2. This results in small global shifts between their velocity maps, which in turn cause differences in the time series that are linear trends.

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 Fig. 2 by a yellow frame, while the zoom over this area is shown in Fig. 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 Fig. 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 Fig. 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_FB, 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.
3. VELOCITY INTER-COMPARISON

This section describes the velocity inter-comparison. 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 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 shown 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% ($\alpha=0.01$), 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.

<table>
<thead>
<tr>
<th>[mm/yr]</th>
<th>Mean</th>
<th>Stdev</th>
<th># PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA</td>
<td>1.21</td>
<td>0.56</td>
<td>27</td>
</tr>
<tr>
<td>TB</td>
<td>1.41</td>
<td>0.43</td>
<td>40</td>
</tr>
<tr>
<td>TC</td>
<td>-0.76</td>
<td>0.55</td>
<td>36</td>
</tr>
<tr>
<td>TD</td>
<td>-0.26</td>
<td>0.33</td>
<td>6</td>
</tr>
</tbody>
</table>
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 minimum and maximum differences, shown in Tables 1 and 2. The minimum values in Table 2 range from -11.46 to -17.04 mm/yr, and the maximum 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).

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.
### 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.

Finally, to get an idea of the spatial distribution of the difference, an example of map of the velocity differences is shown in Fig. 4, which corresponds to the difference between teams A and C. Note that for visualization purposes the PS size has been increased. Fig. 4 corresponds to a portion of the Gardanne mining area. There are dominant greenish colours, which correspond to difference around zero, plus some localized bigger differences, e.g. those in red of about 4-5 mm/yr.

![Figure 4: Zoom over the map of the velocity differences between teams A and C, superposed to the mean SAR amplitude. The zoom covers a portion of the Gardanne mining area.](image-url)
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 6, 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, stdev., min. and max. 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 6 the statistics over the three main parameters are listed by columns. The most significant statistics from Table 6 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 H₀ hypothesis that the mean of the differences is zero. For the six pairs, the H₀ 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.

```markdown
<table>
<thead>
<tr>
<th>Team Pair</th>
<th>Max abs [mm]</th>
<th>Mean [mm]</th>
<th>Stdev [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD-TB</td>
<td>Max</td>
<td>207.67</td>
<td>89.66</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>3.63</td>
<td>-77.02</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>15.20</td>
<td>-0.78</td>
</tr>
<tr>
<td></td>
<td>stdev</td>
<td>7.52</td>
<td>5.84</td>
</tr>
<tr>
<td>TD-TC</td>
<td>Max</td>
<td>196.58</td>
<td>60.12</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>4.47</td>
<td>-92.86</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>16.03</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>stdev</td>
<td>7.51</td>
<td>5.84</td>
</tr>
<tr>
<td>TB-TC</td>
<td>Max</td>
<td>168.32</td>
<td>60.98</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>4.38</td>
<td>-72.52</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>16.33</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>stdev</td>
<td>7.23</td>
<td>5.21</td>
</tr>
</tbody>
</table>

Table 6. 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.

<table>
<thead>
<tr>
<th>Team Pair</th>
<th>Max abs [mm]</th>
<th>Mean [mm]</th>
<th>Stdev [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD-TA</td>
<td>Max</td>
<td>108.03</td>
<td>49.60</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>3.68</td>
<td>-54.51</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>19.00</td>
<td>-0.11</td>
</tr>
<tr>
<td></td>
<td>stdev</td>
<td>8.00</td>
<td>6.54</td>
</tr>
<tr>
<td>TB-TA</td>
<td>Max</td>
<td>155.67</td>
<td>68.63</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>3.07</td>
<td>-82.07</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>13.85</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>stdev</td>
<td>5.90</td>
<td>4.99</td>
</tr>
<tr>
<td>TA-TC</td>
<td>Max</td>
<td>116.33</td>
<td>47.75</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>3.45</td>
<td>-49.61</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>15.71</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>stdev</td>
<td>6.93</td>
<td>5.69</td>
</tr>
</tbody>
</table>
```
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 Fig. 5 and 6. 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 Fig. 6, 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.

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 Fig. 7. In this case the four teams estimate a cumulated deformation up to 200 mm, with rather small differences. 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.

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%). 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.

![Figure 5](image1.png)

**Figure 5.** 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 6](image2.png)

**Figure 6.** 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.

![Figure 7](image3.png)

**Figure 7.** Example of time series where the results of the four teams show good agreement.
We now consider the cases where the time series show important discrepancies. In some cases they reflect the big differences in the deformation velocity estimated by the teams. For instance, as already mentioned above, 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.

5. CONCLUSIONS

This paper 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 km². 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.

This section includes two parts. Firstly, the key inter-comparison results are summarized. Then they 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 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 6). These average values, which range from -0.78 to 0.88 mm, indicate that globally the differences are close to zero.

2) Another 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 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 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.

Below, 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$_{TA}^{TA} -$ mean$_{TB}^{TB}$, 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. 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 paper. 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. Other examples concern the time series. From different plots the results of the PSIC4 project show the same behaviour of the Provence inter-comparison. Again, this fact confirms that the Provence inter-comparison and the PSIC4 project show consistent results.

ACKNOWLEDGMENTS

This work has been funded by the European Space Agency, within the GMES Terrafirma Project, ESRIN/Contract no. 19366/05/I-EC. The PSI data co-registration check was performed using a tool kindly provided by Nico Adam (DLR).

REFERENCES
