Defining the Methods for PSI Validation and Intercomparison

BRGM/RP-55636 -FR
April 2006
Defining the methods for PSI validation and inter-comparison

Marta Agudo - Michele Crosetto (IG)
Daniel Raucoules, Bernard Bourgine, Luc Closset (BRGM)
Chris Bremmer - Hans Veldkamp (TNO)
Douglas Tragheim - Luke Bateson (BGS)
Keywords: PS Interferometry, Coal Mining, Gardanne, levelling

In bibliography, this report should be cited as follows:


© BRGM, 2007. No part of this document may be reproduced without the prior permission of the PSIC4 Validation-Consortium members.
### CONTENTS

1. INTRODUCTION .......................................................................................................................... 6
2. VALIDATION AND INTER-COMPARISON OVERVIEW ............................................................ 7
3. BACKGROUND CONCEPTS ........................................................................................................... 9
   3.1. Expected outcomes of the PSI technique ............................................................................... 9
   3.2. Use of a standard and general procedure ............................................................................. 9
   3.3. Two complementary aspects to be considered in the validation ......................................... 9
   3.4. Discussion on the interpolation of data .................................................................................. 10
4. PRE-PROCESSING STEPS ........................................................................................................... 11
   Task name: Data gathering ............................................................................................................. 12
   Task name: Coordinate transformation ........................................................................................... 15
   Task name: Correct for geocoding errors ...................................................................................... 17
   Task name: Refer the PS deformations to the same area ............................................................... 20
   Task name: Spatial resampling of PS data .................................................................................... 24
   Task name: Temporal resampling of levelling data ...................................................................... 28
5. VALIDATION ACTIVITIES ........................................................................................................... 32
   Task name: Rough check of geocoding quality ............................................................................ 33
   Task name: Time Series Validation ............................................................................................... 36
   Task name: Velocity validation ..................................................................................................... 45
   Task name: Overlying with ancillary data .................................................................................... 47
6. INTER-COMPARISON ACTIVITIES ........................................................................................... 49
   Task name: PSI spatial distribution and densities .......................................................................... 50
   Task name: APS inter-comparison ............................................................................................... 53
   Task name: Geocoding inter-comparison ..................................................................................... 56
   Task name: Inter-comparison of velocity maps ......................................................................... 58

APPENDIX. MEASURES TO ASSESS THE PSI PERFORMANCES ....................................................... 60
   A.1 Plot of the PS vs. levelling time series ...................................................................................... 60
   A.2 Plot of spatial PS profiles, with fixed $\Delta T$, along a levelling line vs. levelling profiles .... 61
   A.3 Mean and standard deviation of the PS vs. levelling time series differences ......................... 62
   A.4 Probability of goodness-of-fit of the PS-time series with respect to the levelling time series ...... 63
   A.5 Mean and standard deviation of the PS vs. levelling velocity differences ............................. 64
   A.6 Plot the velocity differential error as a function of the PS distance ....................................... 65
   A.7 PS density in different areas - procedure ............................................................................. 66
   A.8 Average geocoding errors ....................................................................................................... 74
1. INTRODUCTION

This document is devoted to the definition and description of the validation and inter-comparison methods to be used in the PSIC4 project. This document includes the following main sections:

1. Overview of the involved activities.
2. Description of the pre-processing steps.
3. Description of the validation activities.
4. Description of the inter-comparison activities.

The above three sections include a complete description of each of the pre-processing, validation and inter-comparison tasks of PSIC4, providing the following elements:

- Task name, task responsible, input and output.
- Block diagram of the task.
- Task description, with the definition and illustration of the key features of each task.
- Details of the procedure used to perform each task. The objective of this description is providing all the information which is required to replicate each task.

In the last part of the document an annex provides a list of the main outputs of the PSIC4 validation and inter-comparison activities: the measures used to assess the Persistent Scatterer Interferometry (PSI) performances. In this annex each measure is illustrated by an example.
2. VALIDATION AND INTER-COMPARISON OVERVIEW

The objective of the PSIC4 project is to assess the performances of PSI for land deformation monitoring. The main focus of the project is on the technique output, rather than the technique itself. For this reason the validation team has adopted a “user viewpoint”, assessing the results of black boxes. The key questions of the project are:

1) How well PSI describes land deformation, spatially and temporally?
2) How accurately, and how precisely? These questions are addressed by the validation activities, where the results of each team are compared against reference data.
3) How consistent are the results between teams? This question is addressed by the inter-comparison activities, where the results of the teams are compared to each other.

A global diagram of the main steps in the PSIC4 is shown in Figure 0.1. In this figure three main blocks of activities are included:

- Pre-processing steps, see Figure 0.2.
- Validation activities, see Figure 0.3.
- Inter-comparison activities, see Figure 0.4.

![Figure 0.1: General block diagram of the PSIC4 activities.](image-url)
Figure 0.2: Block diagram of the pre-processing activities.

Figure 0.3: Block diagram of the validation activities.

Figure 0.4: Block diagram of the inter-comparison activities.
3. BACKGROUND CONCEPTS

This section discusses few key background aspects related to the validation and inter-comparison activities of the PSIC4 project.

3.1. Expected outcomes of the PSI technique

Persistent Scatterer Interferometry is a land deformation monitoring technique. As such, it is expected to deliver deformation estimates, which describe spatially and temporally the actual deformation field under analysis, which e.g. can be highly non linear in time. The goal of this document is to describe the methods to assess the capability of PSI to accurately and precisely describe the deformation field analysed in the PSIC4 project. It is important to underline that the teams have worked under “blind conditions”, without any a priori information on the characteristics of the deformation under analysis. This fact should be taken into account during the final assessment of the performances of the teams.

This said, it is worth noting that the validation and inter-comparison activities of PSIC4 do not concern the methodology used by each team. They are only focused on the outputs provided by the teams, checking how they compare with independent reference data.

3.2. Use of a standard and general procedure

For both direct and indirect measurements, the validation procedure most commonly adopted (standard procedure) consists of comparing the output (quantity to be validated) of a given instrument or estimation procedure with an independent estimate of the same quantity, coming from a different source (reference value). Note that the independent estimate usually has a quality intrinsically higher than that of the output to be validated. In the PSIC4 project the same procedure is used, by comparing:

- the indirect land deformation measurements coming from the PSI teams,
- and the measurements (independent estimates) coming from spirit levelling surveys. This is a standard geodetic technique, which guarantees good quality performances in terms of precision and reliability. Over the PSIC4 test side about 1000 monumented points are available, which have been regularly surveyed once or twice per year. The estimated height of each point has an expected standard deviation (1σ) of ± 3-5 mm.

Comment on geoid variation

The levelling measurements are related to the geoid of the area of interest, which is undergoing variations over time due to the mining activity. For the PSIC4 project this has a negligible effect on the validation results. In fact, taking into account the extracted coal volume (e.g. 3 m by 500 m by 2 km) the effect on the geoid (the variation of its ondulation) is well below 0.1 mm (see the T4 report). Compared to the above levelling precision, this effect can be considered negligible.

3.3. Two complementary aspects to be considered in the validation

Two complementary aspects of the PSI measurements will be considered in the validation.

a. The quality of the deformation velocity and time series estimated by PSI, which will be assessed by direct comparison with the reference data coming from levelling. Note that this requires to refer
(put to zero) the PSI datasets with respect to a given reference point or area, which is the same for all datasets, see T7.1.1b.

b. The intrinsic quality of the PSI measurements, taking into account that PSI basically provides “relative measurements”. This can be checked by comparing the differential (point-to-point) deformation measures provided by PSI, which are independent of the chosen reference point, with the differential reference measures. By point-to-point differential measures in this case we mean the deformation differences computed on pairs of Persistent Scatterers (PS). One goal of this task is to derive a diagram that displays the intrinsic precision of PSI vs. the distance between PS.

### 3.4. Discussion on the interpolation of data

In principle, in this work the use of interpolated data should be minimized. However, since PS and levelling data do not coincide in space and in time, two types of data interpolation are needed to carry out some the validation and inter-comparison activities: spatial and temporal interpolation. As discussed below, see the tasks T7.1.2 “Spatial resampling of PS data” and T7.1.3 “Temporal resampling of levelling data”, the data interpolation will based on geostatistical analysis tools. These tools are expected to provide for each interpolated value an estimate of the associated interpolation error.
4. PRE-PROCESSING STEPS

The pre-processing steps include the following tasks:
- Data gathering;
- Coordinate transformation;
- Correct for geocoding errors;
- Refer the PS deformations to the same area;
- Spatial resampling of PS data;
- Temporal resampling of levelling data.

Each task is described below.
## Task name: Data gathering

### Task responsible: BRGM

### Task number: 7.1.1

### Input:
- PSI datasets
- Levelling data
- Orthoimages
- Auxiliary data

### Output:
- PSI datasets
- Levelling data
- Orthoimages
- Auxiliary data

### Block diagram:

```
PSI datasets (1)  Levelling data (2)  Orthoimages  Mining maps, depth of the coal formation
                      |                          |            |
                      |                          |            |
Format consistency check of datasets and format modifications
                      |                          |            |
PSI datasets (1)  Levelling data (2)  Orthoimages  Mining maps, depth of the coal formation
```

### Description:

1. **PSI datasets.** Each team provides for each measured PS the following data (see for details the PSIC4 Phase I deliverable: “Accompanying notes”):
   - WGS84 Geographic coordinates: Latitude, Longitude, Ellipsoidal height; Image coordinates: range, azimuth;
   - Mean PS velocity over the observed period, measured in the radar Line-of-Sight (LOS). In this document, this is named PS velocity;
   - Time series of LOS heights, with one value for each SAR acquisition date. In this document, this is named PS time series;
   - Time series of APS (Atmosphere Phase Screening), with one value for each SAR acquisition date. In this document, this is named APS;
   - The PS phase coherence, with values between [0, 1], which indicates the quality of the PS measurements. In this document, this is called quality index.

2. **Levelling data.** All the ground data were provided by CDF in Lambert III Sud projection. The levelling network consists in 17 traverses covering the mining area, which is 6.5 km wide in the East-West direction and 5 km in the North-South direction.
   - The dataset consists of about 1000 monumented points, which have an average point-to-point distance which ranges between 40-80 m.
   - The temporal frequency of the surveys is annual or biannual.
   - The acquisition procedure has been originally established to achieve a design precision of the
survey of ± 1 cm for the orthometric heights of all measured points. Note that this is a design value, which, as mentioned below, in practice is largely fulfilled by the available dataset.

- The precision of the planimetric positioning of the measured points is ± 10 m.

- The used levelling method is one-way traverse levelling. The levelling is realized by four different teams (Charbonnage de France, CDF; Société du Canal de Provence, SCP; Société Nationale des Chemins de Fer, SNCF; and Pechiney). SCP, SNCF and Pechiney data are not provided with estimation of the closing errors but the same instrument and standard of quality are used by both organisations.

- The data have been provided in a specific format (excel files) suitable for the use of CDF but not adapted to the PSIC-4 project. In particular:
  - separated files were provided for each levelling line, with some format differences,
  - some points were provided without coordinates,
  - some points were not monitored before falling under the influence of the mining works, some other were monitored during all the period of interest,
  - the traverses have different periods of observation.

An example of modified file obtained for the point SNCF 304 (coordinates Lambert III Sud) is given below. The fields BREF, DZcum/33800 and DZCum at t = 33800 have been added as additional characteristics of the points (described in Task 7.1.3).

<table>
<thead>
<tr>
<th>Line</th>
<th>Point</th>
<th>XL3S</th>
<th>YL3S</th>
<th>BREF</th>
<th>DATE</th>
<th>DATE_NUM</th>
<th>DZCum Ini</th>
<th>DZcum /33800</th>
<th>DZcum at t=33800</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>16/11/1992</td>
<td>33924</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>16/12/1992</td>
<td>33954</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>29/03/1993</td>
<td>34057</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>25/05/1993</td>
<td>34114</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>11/08/1993</td>
<td>34192</td>
<td>-6.0</td>
<td>-6.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>26/10/1993</td>
<td>34268</td>
<td>-6.0</td>
<td>-6.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>11/12/1993</td>
<td>34314</td>
<td>-6.0</td>
<td>-6.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>23/02/1994</td>
<td>34388</td>
<td>-6.0</td>
<td>-6.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>10/05/1994</td>
<td>34464</td>
<td>-6.0</td>
<td>-6.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>12/07/1994</td>
<td>34527</td>
<td>-6.0</td>
<td>-6.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>25/01/1995</td>
<td>34724</td>
<td>-4.0</td>
<td>-4.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>06/06/1995</td>
<td>34856</td>
<td>-7.0</td>
<td>-7.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>06/02/1996</td>
<td>35101</td>
<td>-13.0</td>
<td>-13.0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>10/12/1996</td>
<td>35409</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>27/10/1997</td>
<td>35730</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>25/11/1998</td>
<td>36124</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ARBO</td>
<td>SNCF_304</td>
<td>852666</td>
<td>133140</td>
<td>0</td>
<td>14/10/1999</td>
<td>36447</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(3) Other data (like aerial orthophoto and mining map) were gathered from CDF and Regional Geological Survey (SGR).
The formats of these data are:
- orthophoto: IGN compressed format (provided with Mapinfo and ArcGis plugins), in Lambert II étendu.
- mining map: Mapinfo (MIF) format, giving the position of the panels and the dates of the works.
Details of the procedure:
The levelling data were provided by Charbonnage de France (CDF) in 12 Excel files. The following operations were carried out:
1) concatenation of the different files with different formats in order to obtain one file with a common format.
2) consistency check of the dataset (e.g. by removing points without geographic coordinates).
3) removal the data acquired before the period of interest (before 1992).
The other auxiliary data provided by CDF were:
1) mining maps, including the positions of the exploitation fronts at given dates, in MapInfo MIF vector format;
2) depth of the coal formation, in MapInfo MIF vector format. (Other data are available, see T7.2.3). No modification has been applied to these files.
## Task name: Coordinate transformation

<table>
<thead>
<tr>
<th>Task responsible: BRGM</th>
<th>Task number: 7.1.1b1</th>
</tr>
</thead>
</table>

### Input:
- PSI datasets
- Orthoimages

### Output:
- In Lambert III cartographic coordinates\(^{(1)}\):  
  - PSI datasets
  - Orthoimages

### Block diagrams:

1. PSI datasets
2. Coordinate transformation
3. PSI datasets in Lambert III \(^{(1)}\) coordinates
4. Extract a sub-area of Gardanne and coordinate transformation
5. Orthoimage of Gardanne in WGS84
6. Coordinate transformation
7. Orthoimage in Lambert III \(^{(1)}\) coordinates

### Description:

1. **Lambert III cartographic coordinates.** Lambert III Sud is the cartographic coordinates projection used by the official cartography in France for the South of France. The PSIC4 validation and inter-comparison tasks are performed on data referenced to this system.

   The data provided by CDF (levelling and mining maps) have been provided in Lambert III Sud (no conversion was therefore needed). The orthoimage has been provided in Lambert II étendu (official system for all France).

   The conversion of an excerpt of the orthoimage in WGS84 projection has been required by ESA for helping teams in geocoding their data.
**Details of the procedure:**

Import each team dataset into the MapInfo system and project the coordinates of each point from geographic coordinates (Latitude and Longitude in WGS84) into the Lambert III cartographic reference system, by using of the MapInfo ‘Coordinate extractor’ tool (more details on the conversion can be obtained on the MapInfo help at “Coordinate extractor). The new coordinates were stored by inserting two new columns to the file called NL3, EL3.

The orthophoto is directly managed by MapInfo. As an except is needed we propose to use the Envi for extracting an area around the site of Gardanne and to convert it WGS84 (for more details consult the Envi help at resize “Images” and “Projection conversion”).
# Task name: Correct for geocoding errors

<table>
<thead>
<tr>
<th>Task responsible: BRGM</th>
<th>Task number: 7.1.1b2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
<td><strong>Output:</strong></td>
</tr>
<tr>
<td>- PS data geocoded by the teams in Lambert III coordinates</td>
<td>- Corrected geocoded PS data in Lambert III coordinates</td>
</tr>
<tr>
<td>- Orthoimage</td>
<td></td>
</tr>
</tbody>
</table>

## Block diagram:

1. **PS data geocoded by the teams in Lambert III coordinates**
2. **Estimate the average shifts (ΔE, ΔN)**
   - for each team
3. **Corrected geocoded data in Lambert III coordinates**

## Description of PS data:

1. **Average shifts (ΔE, ΔN).** The geocoding of PSI products is usually not accurate, due to the uncertainty associated with two SAR geometric parameters: the time of acquisition of the first SAR line, and the near slant range. This uncertainty results in two geocoding systematic errors, which correspond to two shifts in the cartographic directions: ΔE and ΔN. Despite the availability of an orthoimage during geocoding, the PS data geocoded by the teams show non negligible systematic errors, which have been corrected for, by the validation group, using an orthoimage, see Figures 1 and 2.
Figure 1: Example of systematic geocoding error of PS data: the PS (green squares) do not fall in the exact location of the structure (pipeline).

Figure 2. This diagram shows all the displacement vectors measured for each team during the pre-processing steps. An average vector was produced and used to globally shift the datasets.
Details of the procedure

Plot the points on the IGN orthoimage of the test area. For all the teams a global positioning error on the geocoded results was noticed. It was therefore decided to assess and correct this error for each team. The correction was carried out using MapInfo system.

a) The dataset of Team 4 seemed to be the most suitable reference, though it showed a global shift. We started by shifting the Team 4 dataset by taking the point E=847376.2, N=128952.2 as reference.

b) Team 2, 3, 5, 6, 7, were shifted by taking Team 4 as a reference. The geocoding compensation was double checked visually by using the orthoimage as a reference. In particular, the pipeline close to point E=846911.9, N=128614.8, which seems to be a good source of scatterers for most of the teams, was used for this purpose. Figure 3 shows the position of the PS along the pipeline after correction for Teams 2 to 7.

c) No reliable correspondence for Team 1 and Team 8 were found. Therefore, the correction was not performed for those teams.

The geocoding correction of the PS was validated by BGS in the Task 7.2.1.

<table>
<thead>
<tr>
<th>Team</th>
<th>MOY (m)</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5.07</td>
<td>2.42</td>
</tr>
<tr>
<td>3</td>
<td>2.61</td>
<td>1.90</td>
</tr>
<tr>
<td>4</td>
<td>2.82</td>
<td>2.90</td>
</tr>
<tr>
<td>5</td>
<td>2.76</td>
<td>1.79</td>
</tr>
<tr>
<td>6</td>
<td>2.10</td>
<td>1.60</td>
</tr>
<tr>
<td>7</td>
<td>3.99</td>
<td>3.05</td>
</tr>
</tbody>
</table>

Figure 3: Estimation of mean geocoding error at the reference area. The blue dotted line represents the reference area (metallic pipeline in the orthophoto). The mean and sigma of the distance between each scatterer and the pipeline is indicated in the table.
Task name: Refer the PS deformations to the same area

Task responsible: BRGM  Task number: 7.1.1b3

Input:
- Levelling data
- Orthoimage
- Mining map
- PS deformation velocity maps
- PS deformation time series

Output:
- PS deformation velocity map referred to the same stable area
- PS deformation time series referred to the same stable area

Block diagram:

Description of PS data:

(1) Stable area. PSI provides relative (differential) deformation measurements. For this reason the deformation are usually referred to a reference point, see Figure 4a. In order to perform the validation and inter-comparison activities all PS datasets have been referred to the same stable area, chosen by considering the levelling data and the mining maps.

(2) Compute the mean velocity. In the PSIC4 project, the above mentioned reference is a stable area, where it is assumed to have at least few PS per each team. The mean velocity is computed following the procedure illustrated in Figure 4b.
(3) Subtract the mean velocity from each time series.

An error on the velocity estimation (such those identified after the estimation of the mean velocity on the stable area) produces a linear trend on the time series due to the difference between the actual mean velocity and the estimation (see Figure 5). Correction will consist in removing the linear trend corresponding to the incorrectly estimated velocity on the stable area.

Figure 4a: The differential PS measurements need a reference point or area. All deformations are relative to this reference.

Figure 4b: Scheme of the computation of the mean velocity over the stable area. PS are indicated by green squares.

\[
V_{\text{Mean}} = \frac{\sum_{i=1}^{n} V_{PSi}}{n}
\]

\(V_{PS1}\)

\(V_{PSi}\)

\(V_{PSn}\)

Stable area
Figure 5: For a given PS, an incorrect estimation of the velocity on the stable area will introduce a linear trend on the Time Series. Once the average velocity is estimated on a stable area, this trend can be corrected.

**Details of the procedure:**

Put the data in the same reference.

a) A stable area was selected (i.e. an area where the levelling data show no deformation during the observed period).

Figure 6: Stable area defined around a portion of a levelling traverse containing 5 levelling points far from the mining works. Colour dots show the positions of PS of different teams.
b) The mean LOS velocity for the points in this area was computed for each team. Then, this velocity was subtracted from the velocity of all the points in order to put to a common reference each velocity map.

c) The LOS mean velocity values were also used to subtract linear trends to the time series. Each SAR date is identified by its Julian day. We therefore subtracted to each time series a function \( V(t-t_{\text{ref}}) \), where \( V \) is the LOS velocity in the stable area, \( t \) is the SAR date in Julian days and \( t_{\text{ref}} \) is the selected reference date. Because the provided files were very different (difference in the acquisition dates, field, etc.), we wrote short C programs, specific to each file in order to subtract the linear trend due to the velocity error, see an example in Figure 7.

![Figure 7: Example of a Time Series correction for Team 4.](image-url)
Task name: Spatial resampling of PS data

Task responsible: TNO
Task number: 7.1.2

Input:
- Database of PS results for 8 teams. The results of each Team have been corrected for geocoding errors (Task 7.1.1b2) and referenced to the same stable area (Task 7.1.1b3)

Output:
- Database of estimated PS-InSAR displacement and kriging error at levelling locations at satellite acquisition dates

Block diagram:

Description of PS data:

(1) **Normal-score transformation**: A normal-score transformation is a transformation of the cumulative distribution of original $z$-values into the standard normal distribution of $y$ values, called normal scores. The $n$ original data $z(u_i)$ are ranked in ascending order. The sample cumulative frequency of the datum $z(u_k)$ with rank $k$ is computed. The normal score transform of the $z$-datum with rank $k$ is matched to the $p_k$-quantile of the standard normal cumulative distribution function (cdf).

(2) **Semi-variogram**: The semi-variogram is a measure of spatial continuity of spatially
distributed data and used for weighing data-points in an interpolation scheme known as kriging. It is defined as:

\[
\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha) - z(u_\alpha + h)]^2
\]

Where: \( \gamma(h) \) = semi-variogram value

\( N(h) \) = number of data-points at lag or distance \( h \)

\( z(u_\alpha) \) = measurement at location \( u \)

\( z(u_\alpha + h) \) = measurement at location \( u + h \)

An example of a semi-variogram (experimental values and model) of the normal-score transformed displacements of one of the teams is shown below:

(3) Ordinary Kriging: Ordinary Kriging is a least-squares regression algorithm for estimating the value of a continuous attribute \( z \) at any unsampled location \( u \) using \( z \)-data in a given neighbourhood of \( u \). The objective of Ordinary Kriging is to minimize the estimation variance under the constraint of unbiasedness of the estimator:

Unbiasedness:

\[
E[Z^*-Z_0] = E\left[ \sum_\alpha \lambda_\alpha Z_\alpha - Z_0 \right] = 0
\]

Minimize estimation variance: \( Var[Z^*-Z_0] = Var\left[ \sum_\alpha \lambda_\alpha Z_\alpha - Z_0 \right] \) minimum

Where: \( \lambda_\alpha \) = a properly chosen weight so that the estimation variance is minimized.

This leads to the following system of equations:

\[
\begin{align*}
\sum_\alpha \lambda_\alpha C_{\alpha \beta} + \Lambda &= C_{\beta 0} \\
\sum_\alpha \lambda_\alpha &= 1
\end{align*}
\]

Where \( C \) is the spatial cross-covariance and \( \Lambda \) is a Lagrange multiplier. This system is solved for \( \lambda_\alpha \) and \( \Lambda \) because \( C \) can be inferred from the semi-variogram. Hence, the
weights associated with the measurements $z$ in the neighbourhood of the location $u$ are known.

(4) **Output database**: This is a CSV-file consisting of the following columns:

- $SN^+ = \text{Sample rank number for the estimated levelling points}$
- $Xl3s = \text{x-coordinate of levelling point in Lambert3-Sud projection}$
- $Yl3s = \text{y-coordinate of levelling point in Lambert3-Sud projection}$
- $Day = \text{Julian day}$
- $BREF = \text{code assigned 1 for levelling time series starting before 15 July 1992 and assigned 0 for levelling time series beginning later than 15 July 1992}$
- $Date = \text{date}$
- $Disp = \text{displacement estimated by temporal interpolation of the levelling time series}$
- $Indc = \text{levelling location code}$
- $Type = \text{name of levelling series}$
- $n\text{-team} = \text{number of teams for which a displacement could be estimated at the specified location}$
- $t^*_{\text{cumulative displacement estimate}} = \text{estimated cumulative displacement for team } * \text{ at date } Date \text{ for levelling location } Indc$
- $t^*_{\text{cumulative displacement stdev}} = \text{kriging error for team } * \text{ at date } Date \text{ for levelling location } Indc$

(5) **Kriging error**: The kriging error follows from the minimization condition of the error variance and is defined as:

$$\sigma_{ok}^2 = C_{00} - \sum_{a} \lambda_a C_{a0} - \Lambda$$

**REFERENCES:**


**Details of the procedure:**

Spatial resampling of the PS-data will be 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. 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 results of Task 7.1.1b3 is taken as input. This means that the primary results of the teams
have been corrected for a geocoding error, expressed in a Lambert III-South reference system, and then related to a stable area with a 0-displacement reference.

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 is applied, implying that even if only one data-point is present, an estimation is made of the cumulative displacement at the respective levelling location.

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-coordinate of levelling point in Lambert3-Sud projection
- Yl3s = y-coordinate 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
Task name: Temporal resampling of levelling data

Task responsible: BRGM

Task number: 7.1.3

Input:
- Levelling data
- Acquisition dates

Output:
- Interpolated levelling data and associated errors

Block diagram:

Levelling data

Acquisition dates of SAR images

Refer each levelling line to the same temporal origin as SAR data

Perform a temporal variogram analysis

Temporal interpolation of levelling data at SAR acquisition dates

Interpolated levelling values and the associated errors (kriging error)

Description of PS data:
The description of the key terms is included in the following section.

Details of the procedure:
The pre-processing of the levelling data involves different steps:

1) Refer each levelling line to the same temporal origin

The provided levelling data have different temporal origin. To be consistent with SAR data, for which a zero-time reference for the study was set on 15/07/1992, the cumulated displacements of levelling data were also shifted in order to have a cumulated displacement equal to zero on 15/07/1992 (first common date of the 8 teams for SAR acquisition). For levelling points whose first levelling date is posterior to 15/07/1992, it was assumed that the cumulated displacement before the first date is equal to zero. In practice this assumption was made by the responsible of the levelling network: some points were not monitored before falling under the influence of the mining works. In order to distinguish them from the main set of interpolated data we added to the data set an additional attribute ("BREF") set to 0
when the point was not monitored before the 15/07/1992 and to 1 when it had been monitored before and after this date.

The format of the corrected data is a set of points with the following fields:

- ID of point.
- (E,N) coordinates (Lambert III-South projection).
- Time "coordinate" expressed in number of Julian days (0 = 01/01/1900).
- Cumulated displacement since the reference date (15/07/1992).
- Error attribute, this is an error due to temporal interpolation when interpolating the cumulated displacement on 15/07/1992.

For levelling points that have been surveyed before and after the reference date of 15/07/92 the cumulated displacement on 15/07/1992 was temporally interpolated (by kriging using variogram defined below) between the nearest surveyed dates (before and after 15/07/1992), and the attribute "BREF" set to 1. For levelling points first surveyed after the 15/07/1992, the cumulated displacement on 15/07/1992, the BREF attribute and the error attribute was set to zero.

2) The temporal variogram was computed from levelling data, using all available data (covering a relevant period of time). It was checked (using cross validation techniques), if de-trending is necessary or not, and if kriging standard deviation is well estimated.

The temporal (semi)-variogram was computed using the formula:

$$ \gamma(dt) = \frac{1}{2} \text{Mean}( [DZ(t) – DZ(t+dt)]^2 ) $$

Were DZ(t) and DZ(t+dt) are the cumulated displacements measured at the same levelling point, at times t and t+dt (the "distance" between the two points is in fact a temporal distance).

3) The temporal resampling was done by kriging. For each selected levelling point, the cumulated displacement was interpolated at each SAR date, from the displacements measured at the same point at levelling dates.

4) Results were provided as an Excel file with the following columns:

- Id of point.
- (E,N) coordinate (Lambert III-South).
- Cumulated displacement since the 15/07/1992 reference levelling date for each SAR date.
- BREF attribute.
- error standard deviation due to temporal interpolation when interpolating cumulated displacement at each SAR date.
- Set of n SAR dates, with for each date: interpolated cumulated displacement and the associated kriging standard deviation.
- An assessment of the mean velocity for each point.

Note: the cumulated displacement are relative to the reference date (0 on 15/07/1992).
Figure 8: Temporal variogram of the levelling data set and best fitting curve.

Examples for temporal resampling

Temporal resampling and original levelling data (Line: AXE, point: RNGF_7)
Temporal resampling process

The levelling data (pink squares in the Figure above) are used to interpolate the cumulated displacement at each SAR acquisition date. Results are the blue rhombus. The dashed lines correspond to the interpolated value +/- the kriging standard deviation (i.e. the kriging error) and indicate an estimate of the interpolation error. The interpolation error tends to zero when a SAR date is close to a levelling date. The interpolation error is also smaller when deformation rate is smaller (see for example between June and September 2000 compared to July 1999).

Example of temporal interpolation result

The cumulated displacement (DZ*) at SAR acquisition date t=16/10/1999 is estimated from the 3 following levelling measurements:

- t=27/04/1999 ; DZ= -122.3 mm
- t=15/09/1999 ; DZ= -126.3 mm
- t=18/05/1999 ; DZ= -135.3 mm

The obtained result is DZ* = -127.5 mm, with a kriging standard deviation of 2.6 mm. As the SAR date is close to 15/09/1999, a higher weight is assigned to this levelling point (the two other give the general trend). The estimated DZ* at SAR date t=11/09/1999, using the same levelling data points, is DZ*=-126.2 mm, with a kriging standard deviation of 0.8 mm.
5. VALIDATION ACTIVITIES

The validation activities include the following tasks:
- Rough check of geocoding quality;
- Time Series Validation;
- Velocity validation;
- Overlying with ancillary data.

Each task is described below.
### Task name: Rough check of geocoding quality

<table>
<thead>
<tr>
<th>Input:</th>
<th>Output:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- PS data in Lambert III, geocoded by the teams and then globally repositioned during Task 7.1.1fb&lt;br&gt;- Orthoimage</td>
<td>- Vectors indicating the varying shifts with in the globally repositioned PS dataset</td>
</tr>
</tbody>
</table>

**Block diagram:**

1. **PS data in Lambert III, geocoded by the teams and then globally repositioned during Task 7.1.1fb**
2. **Measure shifts (ΔE, ΔN) in several different locations for each team**
3. **Vectors indicating the varying shifts with in the globally repositioned PS dataset**

**Description of PS data:**

The PSI data have been geocoded by the teams using the orthoimage. These geocoded data were found to be globally shifted, a feature that was assessed and removed in the pre-processing of Task 7.1.1fb. It is now realised that there are smaller residual non-systematic shifts in the data, which require assessing. An example of the varying shift directions and magnitudes is shown in Figure 9.

**Details of the procedure:**

The following is a step-by-step guide 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. The coordinate system of the data frame was set to the Lambert III cartographic reference system.
Figure 9: Data from Team 7 for a rural area. It can be seen that the residual shift (i.e. that remaining after a global shift has been applied during pre-processing) direction and magnitude varies spatially within a teams dataset.

2) The orthorectified, 1 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.
   3a) This data had been pre-processed by BRGM. The preliminary shift to the data was applied.
   3b) 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:
   4a) All teams results were displayed over the orthoimage, each teams points were displayed in a different colour.
   4b) 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 aerial photograph. 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 points' actual location should be. Coordinates for the start and end of the vector were recorded.

7) Differencing the coordinates gave a shift, in meters, for both the E and N directions between the point's observed and actual locations.

8) These shifts were recorded in an Excel document for each location and for each team. For each measurement a screen grab of the location, vector and points was also added a word document along with the coordinates and shifts.

9) E and N shifts for each team were then averaged to gain an overall shift for each team.
### Task name: Time Series Validation

### Task responsible: TNO

**Task Number:** 7.2.2.1

### Input:
- Output database of spatial resampling (Task 7.1.2)
- Output database of temporal resampling (Task 7.1.3)

### Output:
- Time series plots
- Spatial deformation profile
- Mean and the st. deviation of errors
- Non-parametric good-of-fit

### Block diagram:

```
Levelling time series

Plot levelling time series vs. PS time series

Plot levelling vs. PS (fixed Δt) along a levelling line

Statistical analysis

Time series plots
Spatial deformation profile
Root Mean Squared Error
Non-parametric goodness-of-fit
```

### Description of PS data:

1. **Time series plot.** An example of the estimated PS-time series of displacements of a number of teams against the levelling time series at a certain levelling location (in this case SIMI-2506) is shown in the following Figure. The error-bars show the kriging error associated to the ordinary kriging system used to estimate the PS-time series at the given location.
(2) Spatial deformation profile. The following plots show, for a certain levelling line, the cumulative displacement along the levelling transect and, for team 8, the spatially resampled PS results. Levelling profiles are based on the temporally resampled levelling data. The cumulative displacement is colour coded and in mm. PS results are LOSV-corrected and spatially resampled for a 50 metre neighbourhood. The horizontal axis is distance, the vertical axis is Julian date. Along the top horizontal axis the *indc* levelling location codes are shown. The examples are for the SNCF-levelling line. The results of the levelling line and of the various teams is shown.

Cumulative displacement (in mm) of the time-interpolated levelling points along the levelling line SNCF as a function of time (Julian days).
Cumulative displacement (in mm) of the Permanent Scatterers for Team 8 within a 50 metre neighbourhood of the levelling points along levelling line SNCF as a function of time (Julian days).

**Details of the procedure:**

Purpose of this task is to compare time series of displacement at co-located PS- and levelling points and test the hypothesis that both time series are drawn from the same population. The results of Task 7.1.1 “Data gathering”, Task 7.1.2. “Spatial resampling” and Task 7.1.3 “Temporal resampling of levelling data” is taken as input. For this task, the mathematical software package Matlab will be used. The method used is described as follows:

1) For each team the co-located PS-time series and levelling time series are plotted. The PS-time series are derived from the spatial resampling data set. Error-bars for the PS-time series are shown based on the spatial resampling (kriging) error. Along levelling transects the cumulative displacement in time is shown based on the temporal resampling results of the levelling time series and the spatial resampling results of the PS-time series.

2) The RMSE is computed at selected levelling locations such that a representative estimate can be made.

3) A goodness-of-fit test of the PS-time series with respect to the levelling time series is computed based on the non-parametric Kolmogorov-Smirnov-test.

The Kolmogorov-Smirnov goodness-of-fit test (K-S test) is applied for comparing the PS-time series with the levelling time series.

The K-S test tries to determine if two datasets differ significantly. The K-S test has the advantage of making no assumption about the distribution of data, hence it is non-parametric and distribution free.

The Kolmogorov-Smirnov (K-S) test is based on the empirical cumulative distribution function (ECDF). Given $N$ data points $x_1, x_2, ..., x_N$, which are ordered from smallest to largest value, the ECDF $S_N(x)$ is defined as
\[ S_N(x) = \frac{\text{number of samples } \leq x_i}{N} \] (1)

This is a step function that increases by \(1/N\) at the value of each ordered data point.

While comparing two datasets \(X_1\) and \(X_2\) consisting of respectively \(N_1\) and \(N_2\) ascendingly sorted data points, the ECDF for these two datasets are defined as:

\[ S_{N_1}(x) = \frac{\text{number of samples } \in X_1 \leq x_i}{N} \] (2)

\[ S_{N_2}(x) = \frac{\text{number of samples } \in X_2 \leq x_i}{N} \] (3)

where \(x_i \in X\), \(X = X_1 \cup X_2\), and \(N\) is the number of data points in \(X\), which are ascendingly sorted. \(S_{N_1}(x)\) is the proportion of \(X_1\) values less than or equal to \(x_i\) and \(S_{N_2}(x)\) is the proportion of \(X_2\) values less than or equal to \(x_i\).

The Kolmogorov-Smirnov statistics (K-S statistics) \(D_o\) is defined as the maximum value of the absolute difference between two cumulative distribution functions. Thus, for comparing two ECDF \(S_{N_1}(x)\) and \(S_{N_2}(x)\) the K-S statistics is

\[ D_o = \max_{-\infty < x < \infty} \left| S_{N_1}(x) - S_{N_2}(x) \right| \] (4)

Figure 10 gives an example of ECDF and how the K-S statistics is determined.

The function that enters into the calculation of the significance level in a K-S test is given by

\[ Q_{KS}(\lambda) = 2 \sum_{k=1}^{n} (-1)^{k-1} \exp(-2k^2\lambda^2) \] (5)

which is a monotonic function with the limiting values \(Q_{KS}(0) = 1\) and \(Q_{KS}(\infty) = 0\).

In terms of this function, the significant level of an observed value of \(D\) – as a disproof of the null hypothesis that the distribution are the same – is given approximately (von Mises, 1964) by the formula

\[ P(D > D_o) = Q_{KS} \left( \left\lceil \sqrt{N_e} + 0.12 + \frac{0.11}{N_e} \right\rceil D_o \right) \] (6)

where \(N_e\) is the effective number of data points and is defined as follows:

\[ N_e = \frac{N_1 N_2}{N_1 + N_2} \] (7)

When \(P(D > D_0)\) is close to 0, the two datasets are probably from different distribution. Usually, the hypothesis is rejected if the test is significant at the 5% level.
Figure 10: Example of empirical cumulative distribution functions made from two different datasets.

Another parameter that can be used to quickly evaluate the quality of the relationship between the two datasets is the root mean square 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 represents the typical size of discrepancy between the two datasets, 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. The squared difference term places more weight on large discrepancies.

Example 1 – Datasets for point ARBO_12_901 (x = 850206, y = 133577).

We consider the temporally interpolated levelling data and the spatially interpolated PS data for point ARBO_12_901. The latter is preprocessed by Team 8. Figure 11 presents these two datasets.
Figure 11: Temporally interpolated levelling data and the spatially interpolated PS data, preprocessed by Team 8 for point ARBO_12_901 (3 points at average distance 39 meters and average coherence 0.51).

The K-S statistics is $D_0 = 0.1035$ and the K-S probability $p = 0.9861$, indicating a good fit between the two datasets. The ECDF of these two datasets are presented in Figure 12.

Figure 12: Empirical cumulative distribution functions made from the time-interpolated levelling data and spatial-interpolated PS data of Team 8 for point ARBO_12_901.

The RMSE for these two datasets is calculated as 8.60 mm.
Example 2 – Datasets for point SNCF_038 (x = 8525885, y = 134289)

For this point PS datasets preprocessed by Team 3, 4, 5, 7 and 8 are available.

Figure 13 presents the temporally interpolated levelling data and the spatially interpolated PS datasets.

![Figure 13: Temporally interpolated levelling data and the spatially interpolated PS datasets for point SCNF_038.](image)

Table 1 and 2 present the results of respectively the KS test and the RMSE calculations.

<table>
<thead>
<tr>
<th>Team</th>
<th>KS probability</th>
<th>KS statistics</th>
<th>Hypothesis Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4.99E-6</td>
<td>0.4616</td>
<td>rejected</td>
</tr>
<tr>
<td>4</td>
<td>6.40E-7</td>
<td>0.4970</td>
<td>rejected</td>
</tr>
<tr>
<td>5</td>
<td>1.21E-10</td>
<td>0.6207</td>
<td>rejected</td>
</tr>
<tr>
<td>7</td>
<td>0.6752</td>
<td>0.1323</td>
<td>accepted</td>
</tr>
<tr>
<td>8</td>
<td>2.01E-13</td>
<td>0.8103</td>
<td>rejected</td>
</tr>
</tbody>
</table>

**Table 1**: Summary of KS test results for datasets of point SNCF_038.

<table>
<thead>
<tr>
<th>Team</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>27.52</td>
</tr>
<tr>
<td>4</td>
<td>15.12</td>
</tr>
<tr>
<td>5</td>
<td>17.73</td>
</tr>
<tr>
<td>7</td>
<td>5.7137</td>
</tr>
<tr>
<td>8</td>
<td>27.52</td>
</tr>
</tbody>
</table>

**Table 2**: RMSE values for datasets of point SNCF_038.

It appears from the results that the dataset preprocessed by Team 7 have a good fit with the levelling data. The null hypothesis that other datasets are of the same distribution as the levelling data, is rejected.
Figure 14: Empirical cumulative distribution functions made from the time-interpolated levelling data and data spatial-interpolated from the PS data preprocessed by Team 7 for point SNCF_038.

Figure 14 presents the ECDF of the levelling data (left) and the dataset that is spatially interpolated from the PS data preprocessed by Team 7 (right).

Figure 15: Empirical cumulative distribution functions made from the time-interpolated levelling data and data spatial-interpolated from the PS data preprocessed by Team 8 for point SNCF_038.

Figure 15 presents the ECDF of the levelling data (left) and the dataset that is spatially interpolated from the PS data preprocessed by Team 8 (right).
REFERENCES:


**Task name: Velocity validation**

**Task responsible:** BRGM  
**Task Number:** 7.2.2.2

**Input:**
- Interpolated PS LOS deformation velocities
- Levelling data

**Output:**
- Mean and standard deviation of the PS vs. levelling velocity differences (errors)
- Plot of differential velocity error as function of distance

**Block diagram:**

1. Interpolated PS deformation velocities
2. Levelling data
3. Transform LOS velocities into vertical velocities
4. Compute the deformation velocity values from the levelling data
5. Compute the mean (bias) and standard deviation of the PS vs. levelling differences (errors)
6. Mean and standard deviation of the PS vs. levelling velocity differences (errors)
7. Map of the velocity errors
8. Compute the differential velocity error for pairs of PS/levelling points
9. Plot the differential velocity error as function of distance

**Description of PS data:**
The description of the key terms is included in the following section.
### Details of the procedure:

Input data: ID of selected levelling points, E, N of this point; spatially interpolated PS velocity value; interpolation error (results of Task 7.1.2). The procedure involves the following steps:

1) The velocity was computed for each of the selected levelling points using a linear regression (following the same method to compute the velocity used for the SAR data).

2) The difference \( \Delta_1 = VPS(i) - VLEV(i) \) was computed for each team, where \( VPS \) is the velocity of PS data, \( VLEV \) the velocity of levelling data, and \( i \) the number of the levelling point. The mean and standard deviation of \( \Delta_1 \) were analysed, and the scatter plots estimated. This step was done directly under ISATIS software. Output: standard deviation, mean, scatter plot for each team.

3) The data is included in a file with the following format:

   Levelling point i; Xi coordinate; Yi coordinate; Velocity_PSi, Velocity_Levelling_I

4) Compute for each line (using a C procedure) of the file:
   - the coordinate differences, the corresponding distance and the differential velocity error
     \( \Delta_2 = (VPS(i) - VPS(j)) - (VLEV(i)-VLEV(j)) \) respect all the other lines of the file.
   - Write in a file distance\_i\_j and Delta2\_i\_j

   Output: file containing differential error and distances, format: Distance; Delta2.

5) Results were displayed on a graph \( \Delta_2 = f(\text{inter-distance } i-j) \).

6) The output is a table giving mean and standard deviation of \( \Delta_1 \) values, for all the teams, and 8 graphs, 1 for each team, showing \( \Delta_2 = f(\text{inter-distance}) \).
## Task name: Overlying with ancillary data

<table>
<thead>
<tr>
<th>Task responsible: BGS</th>
<th>Task number: 7.2.3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
<td><strong>Output:</strong></td>
</tr>
<tr>
<td>- Mining maps</td>
<td>- Spatial relationships between PS deformation velocities and ancillary data</td>
</tr>
<tr>
<td>- Dates when panels were completed</td>
<td></td>
</tr>
<tr>
<td>- Structure contour maps of coal seams</td>
<td></td>
</tr>
<tr>
<td>- Faults and amount of fault displacements</td>
<td></td>
</tr>
<tr>
<td>- Total thickness of coal extracted</td>
<td></td>
</tr>
<tr>
<td>- Water levels</td>
<td></td>
</tr>
<tr>
<td>- Mine water pumping history</td>
<td></td>
</tr>
</tbody>
</table>

## Block diagram:

- Mining maps, total thickness of coal extracted, water levels, mine water pumping history, etc.
- PS deformation velocity maps
- Data analysis
- Spatial relationships between PS deformation and ancillary data

## Description:

The description of the key terms is included in the following section.

## Details of the procedure:

1) The same ArcGIS 9.1 project was used for this task as used in task 7.2.1.
2) The following ancillary data were then converted from MapInfo format to ESRI shapefiles, projected to the Lambert III coordinate system and loaded into the project:
   2a) Geological structures and amount of displacement.
   2b) Dates of advancement.
   2c) Exploitation panels with dates.
   2d) Thickness of coal extracted.
   2e) Structure contours of coal seams.
3) Spatial relationships were then found between the PS points and the mining information. The exploitation panel data was used to select the PS points that fell within the panel area. Data for these points were extracted from the main data table and stored in a new table. This was done for each team.

4) Histories for the PS points within the mined areas were examined, especially with reference to the date that the panel was mined.
   4a) Plots of date and ground movement were produced from the point histories contained in the new tables created in step 3.
   4b) On the plots the date of mining completion was marked therefore allowing a comparison between mining completion and PS movement.

5) Point histories adjacent to the mined areas were also examined to try and establish the area of influence of the mining on ground motion.
   5a) Mining panel data was buffered by different distances (multiples of 100 m). For each buffer and each team PS data were extracted. This produced rings of PS points at different distances from the mined panel. These were extracted from the main table.
   5b) Examination of plots of point motion within these rings allows for a relationship between ground motion and distance from mining to be examined.

6) Point histories were also related to thickness of coal extracted. This was completed for each team using the plots created in step 4a above. The observed ground motion was examined in the light of the thickness of coal extracted in that particular panel.

7) Spatial relationships were and other findings from this task were presented in a report.
6. INTER-COMPARISON ACTIVITIES

The inter-comparison activities include the following tasks:
- PSI spatial distribution and densities;
- APS inter-comparison;
- Geocoding inter-comparison;
- Inter-comparison of velocity maps.

Each task is described below.
**Task name: PSI spatial distribution and densities**

<table>
<thead>
<tr>
<th>Task responsible: BGS</th>
<th>Task number: 7.3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
<td><strong>Output:</strong></td>
</tr>
<tr>
<td>- PS deformation velocity maps</td>
<td>- Plots of overall PS density for each team</td>
</tr>
<tr>
<td>- Differential InSAR results</td>
<td>- Density value (Number of PS per square kilometre) overall, and for rural, urban and unstable areas</td>
</tr>
<tr>
<td>- Orthoimage</td>
<td>- Visual plots of density in urban and unstable areas</td>
</tr>
</tbody>
</table>

**Block diagram:**

1. Orthoimage
2. Classify the area in stable/unstable areas
3. PS deformation velocity maps
4. Study the PS spatial distribution and density
5. Overall density plots
6. Density values for different areas
7. Density plots for urban and unstable areas
**Description and details of the procedure:**

In the first instance the density of the PS results was examined on a dataset-wide approach for each of the eight teams. This approach allowed us to gain an appreciation of areas of high and low densities and see the spatial distribution of these densities. Density was then examined on a team-by-team basis for a smaller area, an area where all teams had PS points.

**Overall density plots for each team**

Density was calculated using hexagonal binning. In other words the number of PS points per hexagonal area was calculated. Hexagonal areas were used for this for the following reasons:

1. Hexagons have symmetry of nearest neighbours, which is lacking in square bins.
2. Hexagons are the maximum number of sides a polygon can have for a regular tessellation of the plane, so in terms of packing a hexagon is 13% more efficient for covering the plane than squares. This property translates into better sampling efficiency at least for elliptical shapes.
3. Lastly hexagons are visually less biased for displaying densities than other regular tessellations. For instance with squares our eyes are drawn to the horizontal and vertical lines of the grid.

To plot the hexagonal density the statistical package S-plus was used. The ‘hexbin’ function was used in two manners:

1) Using the same X and Y-axis values. This allowed a direct comparison on the density and spatial coverage of the teams datasets,
2) Using different X and Y-axis values. This allowed the area of presentation to be maximised for each team.

Greyscale plots of the results allow comparison between teams.

**Density for different areas**

Density was then examined more closely with respect to the land cover type. Firstly an overall figure for density (defined here as number of PS points per kilometre squared) was computed. In order to ensure comparability between the teams the same spatial area had to be examined for each team.

A rectangle was digitised in the ArcGIS project, this rectangle defined an area in which all teams had results. To calculate the areas the following was done:

1) Use the ‘Calculate Areas’ tool in the ‘Utilities’ section of the ‘Spatial Statistic Tools’ within ArcToolbox to find the area of the digitised rectangle.
2) Define a table in MS Excel such as below to calculate the densities.
3) Use the ‘selection’ tool in ArcGIS to select only the PS from each team that fall within the digitised rectangle.
4) Save this selection as a new layer.
5) Open the attribute table of this new layer to find the number of points selected and enter this into the MS Excel table under ‘No. PS in rectangle’. The ‘Density of rectangle’ column is therefore completed for each team.
6) Digitise urban area polygons from the orthorectified aerial photograph and over lay these on the PS data.
7) Use the ‘Calculate Areas’ tool in the ‘Utilities’ section of the ArcToolbox ‘Spatial Statistic Tools’ to find the area of the digitised urban areas. Enter value into ‘Area Urban’ column.

8) Use the ‘selection’ tool in ArcGIS to select only the PS from each team that fall within the urban areas. Save this selection as a new layer.

9) Open the attribute table of this new layer to find the number of points selected and enter this into the MS Excel table under ‘No. PS in urban areas’. The ‘Density of urban’ column is therefore completed for each team.

10) The rural area is found by subtracting the urban area from the rectangular area.

11) DiffInSAR results were used to digitise the unstable regions within the rectangle.

12) The same procedure (6-9) was used to find the PS density for each team within this area.

Visual assessment of the density

To gain a visual appreciation of how land cover types affect the density of the PS points it was necessary to produce a georeferenced density grid for each team. This density grid could then be overlaid with the urban areas identified above.

To produce the density grids, the ‘Spatial Analyst’ extension in ArcGIS was used. One of the options in this extension is the ‘Density’ option. The following parameters are required:

1. Input data. This is the PS point dataset.
2. Population field. This is left at ‘None’ we could enter a different field from the table if we wanted to examine the density of different attributes. This is not relevant to this situation.
3. Density Type, can be set to ‘Kernel’ or ‘Simple’. Simple was used, as this is a measure of points per unit area whereas kernel is a weighted measure of points within a mathematically defined surface.
4. Search Radius, was set to 100, this is the distance to search for points from the centre of each cell in the output.
5. Area units. This was set to Square meters.
6. Output cell size. This was set to 100 meters. 100-meter cells were chosen as these allowed a good compromise between detail of results and processing speed.

The result is a gridded dataset, with a value every 100m. Since we were instructed not to make comparisons between teams it was decided to display each teams data with a colour ramp applied with a 2 standard deviation stretch, Figure 28. This display method highlighted which part of a team’s dataset had the highest density. These density grids were displayed with the urban areas polygons allowing a comparison to be made between the two datasets.
**Task name:** APS inter-comparison

<table>
<thead>
<tr>
<th>Task responsible: BRGM</th>
<th>Task number: 7.3.2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
<td><strong>Output:</strong></td>
</tr>
<tr>
<td>- APS of each SAR image</td>
<td>- APS plots and statistics</td>
</tr>
</tbody>
</table>

**Block diagram:**

1. APS of each SAR image
2. Choose 3 APS for team
3. Resample the data to a 50 by 50 m grid
4. Comparison between teams by pairs: Scatter plots¹ for the selected dates; correlation values² sets for the selected dates; mean and standard deviation of the APS differences

**Description:**

1. **Scatter plot.** The result from one team will be plotted against other team. Ideally, if the results are consistent, the points can be fitted by a straight line with a slope of 100%. An example of scatter plot is shown in the next page.

2. **Correlation matrix:** this matrix represents the set of the correlations between all the possible pairs (e.g. correlations between Ti and Tj i≠j). The correlation is given by:

\[
\rho_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{(n - 1)s_x s_y}
\]

Where \(\bar{x}\) and \(\bar{y}\) are the mean values for \(x_i\) and \(y_i\) and \(s_x\) and \(s_y\) their standard deviations.
Example of scatter plot and correlation value of the APS for the day 26/10/96 (respect to 15/16/92) for Teams 1 and 3. This plot was produced with the software ISATIS.

Details of the procedure:
This task involved the following steps:

1) Selection of 3 common dates (in 1992, 1998, 2004), and of an area covered by the 8 datasets (used software: MapInfo).

Prior to perform the velocity and APS validation is necessary to resample the data on a 50 by 50 m grid. This has been carried out with the ISATIS software. This cell size is consistent with the radius used for the Task 7.1.2. This ISATIS file contains for each cell:
- LOS velocity of each Team (averaged on 50m).
- APS (selected dates 27/10/96, 29/01/00 respect to 15/07/92).
- 3x28 differences Team (i) – Team (j) for LOS velocity and APS. Corresponding to all possible combinations for (28 velocity differences, 2x28 APS differences).

We can assume that the different teams used different reference dates for the APS. We therefore subtracted the first APS from the 2 other dates in order to obtain APS respect the same reference date. The 3 dates have been selected in order to be regularly distributed on a period common to all teams.

This format allows us to easily assess the standard deviation of the differences and the correlation values which are the indicator we have to provide. In addition scatter plots can be derived in order to represent graphically the correlation between results. The number of points used for each comparison is also estimated, as the density and distributions are different from team to team.

The area has been chosen in order to be the largest covered by all the teams. The limits of the grid are expressed in Geographic (Lambert III South), from 846300 to 855800 East and from 126900 to 138150 North.
2) Selection of cells which contain at least 1 PS for each team.

3) For each cell the 2 APS (respect to the reference date) values will be averaged into the cell (used software: ISATIS).

The APS is not the main quality indicator for the PS, but it is an intermediary product needed for the estimation of the deformation, as the interferometric phase is the sum of the APS and the deformation component. Therefore, for obtaining accurate deformation results, the PSI algorithms should correctly estimate the APS. The relative performances of the APS estimation ability of each technique will be therefore tested.

We have to notice that the APS represent a spatially “smooth” dataset: therefore a 50 by 50 m sampling is by far better than the sampling really needed (hundreds of metres). We keep this sampling in order to have a grid common to APS and velocity validation. That has no incidence on the validity of the comparison.

4) Assessment of the correlation matrix and standard deviation of pair differences (software GDM and ISATIS). Under ISATIS software the output grid file is derived directly:

- The correlations between APS for dates 27/10/96, 29/01/00 corrected respect to 15/07/92.
- The standard deviation of the differences between APS.

5) OUTPUT: correlations values, standard deviations of differences, scatter plots, factual analysis. Note: this task is only applicable to the teams that have provided the APS estimations.
Task name: Geocoding inter-comparison

<table>
<thead>
<tr>
<th>Task responsible: TNO</th>
<th>Task number: 7.3.3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong></td>
<td><strong>Output:</strong></td>
</tr>
<tr>
<td>- PS-locations in radar coordinates and Lambert-III-coordinates (task 7.1.1b1)</td>
<td>A matrix with for each couple of teams:</td>
</tr>
<tr>
<td></td>
<td>- Average geocoding difference</td>
</tr>
<tr>
<td></td>
<td>- Standard deviation of the geocoding difference</td>
</tr>
</tbody>
</table>

**Block diagram:**

```
PS-locations in radar coordinates and Lambert-III-coordinates

Analysis

Average geocoding difference

Standard deviation of the geocoding difference
```

**Description:**

1. **Average Geocoding difference**: Average difference in geocoding between two teams based on a comparison of the spatial coordinates assigned to the same radar grid cell.

2. **Standard deviation of the geocoding difference**: Standard deviation of the distribution obtained under ‘Average geocoding difference’.

**Details of the procedure:**

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’ and the standard deviation as the ‘standard deviation of the geocoding difference’. This procedure is followed for each set of two teams and results in the following table of average and standard deviation of the geocoding differences in metre:
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>40.9</td>
<td>45.9</td>
<td>48.6</td>
<td>32.5</td>
<td>136.8</td>
<td>41.8</td>
<td>70.1</td>
</tr>
<tr>
<td>2</td>
<td>40.9</td>
<td>-</td>
<td>17.9</td>
<td>16.9</td>
<td>16.7</td>
<td>112</td>
<td>21</td>
<td>44.6</td>
</tr>
<tr>
<td>3</td>
<td>45.9</td>
<td>17.9</td>
<td>-</td>
<td>11.8</td>
<td>8.2</td>
<td>97.7</td>
<td>21.9</td>
<td>39.3</td>
</tr>
<tr>
<td>4</td>
<td>48.6</td>
<td>16.9</td>
<td>11.8</td>
<td>-</td>
<td>17.6</td>
<td>95.5</td>
<td>18.7</td>
<td>43.5</td>
</tr>
<tr>
<td>5</td>
<td>32.5</td>
<td>16.7</td>
<td>8.2</td>
<td>17.6</td>
<td>-</td>
<td>112.5</td>
<td>20.8</td>
<td>47.1</td>
</tr>
<tr>
<td>6</td>
<td>136.8</td>
<td>112</td>
<td>97.7</td>
<td>95.5</td>
<td>112.5</td>
<td>-</td>
<td>116</td>
<td>75.4</td>
</tr>
<tr>
<td>7</td>
<td>41.8</td>
<td>21</td>
<td>21.9</td>
<td>18.7</td>
<td>20.8</td>
<td>116</td>
<td>-</td>
<td>44.4</td>
</tr>
<tr>
<td>8</td>
<td>70.1</td>
<td>44.6</td>
<td>39.3</td>
<td>43.5</td>
<td>47.1</td>
<td>75.4</td>
<td>44.4</td>
<td>-</td>
</tr>
</tbody>
</table>

Average geocoding difference (in m).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>23.5</td>
<td>20</td>
<td>24.2</td>
<td>18.1</td>
<td>27</td>
<td>24.7</td>
<td>21.1</td>
</tr>
<tr>
<td>2</td>
<td>23.5</td>
<td>-</td>
<td>9.6</td>
<td>14.3</td>
<td>12.3</td>
<td>21.3</td>
<td>14.8</td>
<td>18.3</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>9.6</td>
<td>-</td>
<td>5</td>
<td>8.3</td>
<td>19.8</td>
<td>3.2</td>
<td>13.2</td>
</tr>
<tr>
<td>4</td>
<td>24.2</td>
<td>14.3</td>
<td>5</td>
<td>-</td>
<td>15.3</td>
<td>19.3</td>
<td>7.3</td>
<td>17.8</td>
</tr>
<tr>
<td>5</td>
<td>18.1</td>
<td>12.3</td>
<td>8.3</td>
<td>15.3</td>
<td>-</td>
<td>19.8</td>
<td>6.8</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>21.3</td>
<td>19.8</td>
<td>19.3</td>
<td>19.8</td>
<td>-</td>
<td>23.4</td>
<td>23</td>
</tr>
<tr>
<td>7</td>
<td>24.7</td>
<td>14.8</td>
<td>3.2</td>
<td>7.3</td>
<td>6.8</td>
<td>23.4</td>
<td>-</td>
<td>21.4</td>
</tr>
<tr>
<td>8</td>
<td>21.1</td>
<td>18.3</td>
<td>13.2</td>
<td>17.8</td>
<td>14</td>
<td>23</td>
<td>21.4</td>
<td>-</td>
</tr>
</tbody>
</table>

Standard deviation of geocoding difference (in m).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>84</td>
<td>368</td>
<td>405</td>
<td>208</td>
<td>109</td>
<td>608</td>
<td>219</td>
</tr>
<tr>
<td>2</td>
<td>84</td>
<td>-</td>
<td>12348</td>
<td>4075</td>
<td>1547</td>
<td>381</td>
<td>11432</td>
<td>2649</td>
</tr>
<tr>
<td>3</td>
<td>368</td>
<td>12348</td>
<td>-</td>
<td>14573</td>
<td>10009</td>
<td>1758</td>
<td>39293</td>
<td>6811</td>
</tr>
<tr>
<td>4</td>
<td>405</td>
<td>4075</td>
<td>14573</td>
<td>-</td>
<td>2985</td>
<td>1499</td>
<td>21793</td>
<td>3231</td>
</tr>
<tr>
<td>5</td>
<td>208</td>
<td>1547</td>
<td>10009</td>
<td>2985</td>
<td>-</td>
<td>1057</td>
<td>10972</td>
<td>3113</td>
</tr>
<tr>
<td>6</td>
<td>109</td>
<td>381</td>
<td>1758</td>
<td>1499</td>
<td>1057</td>
<td>-</td>
<td>2113</td>
<td>781</td>
</tr>
<tr>
<td>7</td>
<td>608</td>
<td>11432</td>
<td>39293</td>
<td>21793</td>
<td>10972</td>
<td>2113</td>
<td>-</td>
<td>7725</td>
</tr>
<tr>
<td>8</td>
<td>219</td>
<td>2649</td>
<td>6811</td>
<td>3231</td>
<td>3113</td>
<td>781</td>
<td>-</td>
<td>7725</td>
</tr>
</tbody>
</table>

Number of pairs of radar targets which have been compared per tie.
Task name: Inter-comparison of velocity maps

Task responsible: BRGM  Task number: 7.3.4

Input:
- PS data

Output:
- Comparison between teams by pairs

Block diagram:

PS data

Classify PS data in a regular grid, one for each team

Identify cells with at least 1 PS for each of the teams

Statistical analysis

Inter-comparison based on the statistical analysis

Comparison between teams by pairs:
Scatter plots for the velocities results; correlation values sets; mean and standard deviation of the velocity differences

Description:
The indicators used in this task are the same defined in 7.3.2 (APS inter-comparison). An example of scatter plot is shown below.
Scatter plot and correlation value between the LOS velocities obtained by 2 teams.

**Details of the procedure:**

This task involved the following steps:

1) Selection of an area covered by the 8 data sets (used software: MapInfo).
2) Resampling to a 50 by 50 m grid. Prior to velocity and velocity validation it is necessary to resampled the data on a 50 by 50 m grid, carried out with the ISATIS software. This cell size is consistent with the radius used for the task 7.1.2. We use the same ISATIS file that used in APS comparison. It contains for each cell:
   - LOS velocity of each Team (averaged on 50 by 50 m).
   - APS (selected dates 27/10/96, 29/01/00 respect to 15/07/92).
   - 3x28 differences Team (i) – Team (j) for LOS and APS. Corresponding to all the possible combinations for dates (28 velocity differences, 2x28 APS differences).
3) Selection of cells which contain at least 1 PS for each team.
4) For each point the velocities values will be averaged into each cell (used software: GDM)
5) Assessment of the correlation matrix and standard deviation of pair differences (software GDM And ISATIS). Under ISATIS software the output grid file is derived directly:
   - The mean of differences between velocities.
   - The standard deviation of the differences between velocities.
   - The correlations between velocity values provided by each team.

An additional test on the teams having shown the best performances and a sufficiently correct geocoding with smaller cell sizes (25m and 12.5m) in order to assess the influence of this parameter on the correlation and standard deviation values. OUTPUT: correlations values, standard deviations of differences, scatter plots, factual analysis.

OUTPUT: correlations values, st. deviations of differences, scatter plots, factual analysis.

The area has been chosen in order to be the largest covered by all the teams. These are area limits in Lambert III South: 846300-855800 East, 126900- 138150 North.
APPENDIX. MEASURES TO ASSESS THE PSI PERFORMANCES

A.1  Plot of the PS vs. levelling time series

The plot of the PS vs. levelling line shows for each levelling point the time series of the levelling point and time series of the LOSV-corrected cumulative displacement of Permanent Scatterers within a neighbourhood of 50 metre of the levelling point. Cumulative displacement is in millimetre and set to zero on the reference date of 15 July 1992 (first InSAR acquisition date).
A.2 Plot of spatial PS profiles, with fixed $\Delta T$, along a levelling line vs. levelling profiles

These plots show, for a certain levelling line, the cumulative displacement along the levelling transect and, for each team, the spatially resampled PS results. Levelling profiles are based on the temporally resampled levelling data. The cumulative displacement is colour coded and in millimetre. PS results are LOSV-corrected and spatially resampled for a 50 m neighbourhood. The horizontal axis is distance, the vertical axis is Julian date. The examples are for the SNCF-levelling line and Team 8.

Cumulative displacement (in mm) of the levelling points along levelling line SNCF as a function of time (Julian days).

Cumulative displacement (in mm) of the Permanent Scatterers for Team 8 within a 50 metre neighbourhood of the levelling points along levelling line SNCF as a function of time (Julian days).
A.3 Mean and standard deviation of the PS vs. levelling time series differences

Defining, for each SAR image acquisition time, $d$ as the differences between the deformation measured by the PS and the one measured by levelling, two key measures to assess the PSI performances are given by the mean and the standard deviation of the vector $d_i$, where $i=1,..., N$ the number of SAR acquisitions.
A.4 Probability of goodness-of-fit of the PS-time series with respect to the levelling time series

The Kolomogorov-Smirnov goodness-of-fit test (K-S test) is applied for comparing the PS-time series with the levelling time series.

As an example, we consider the temporally interpolated levelling data and the spatially interpolated PS data for point ARBO_12_901. The following Figure presents these two datasets. The K-S statistics is:

\[ D_0 = 0.1035 \]

and the K-S probability is:

\[ p = 0.9861, \]

which indicates a good fit between the two datasets.
A.5 Mean and standard deviation of the PS vs. levelling velocity differences

For a given levelling point (indicated by \( i \) in the formula) position, the spatial resampling of PS-InSAR data and temporal resampling of levelling allows to associate to this position a velocity value derived from PS-InSAR (\( V_{ps} \)) and a velocity derived from levelling (\( V_{lev} \)). Levelling velocity is computed by applying a linear regression to the temporal series measured by levelling in order to be comparable with PSINSAR velocity. The velocity (\( V_{lev} \)) corresponds to the slope of the regression.

- Mean of the differences:

\[
Md = \frac{\sum_{i} (V_{ps_i} - V_{lev_i})}{N}
\]

- Standard deviation of the differences:

\[
Sd = \sqrt{\frac{\sum_{i} (V_{ps_i} - V_{lev_i} - Md)^2}{N}}
\]

If the standard deviation is an indicator on the precision of the measure, the mean will provide information on possible biases on the velocity assessment. Both values are needed to fully characterise the error on the measure.
A.6 Plot the velocity differential error as a function of the PS distance

For two different points (selected levelling points), we will compute the functional:

\[
C_{i,j} = (\delta d_{ps}(i, j) - \delta d_{lev}(i, j))^2
\]

Where:

\[
\delta d_{ps}(i, j) = V_{ps}(j) - V_{ps}(i)
\]

\[
\delta d_{lev}(i, j) = V_{lev}(j) - V_{lev}(i)
\]

The square root of \(C_{i,j}\) will be plotted against the distance between the points \(i\) and \(j\).

This measure will provide information about the differences of spatial behaviour between PSI and levelling showing scale dependent biases. In particular, uncorrected atmospheric effects could have a such spatial behaviour.
A.7 PS density in different areas - procedure

MEASURING AND ASSESSING PS DENSITY IN DIFFERENT AREAS
In the first instance the density of the PS results was examined on a dataset-wide approach for each of the eight teams. This approach allowed us to gain an appreciation of areas of high and low densities and see the spatial distribution of these densities. Density was then examined on a team-by-team basis for a smaller area, an area where all teams had PS points.

HEXAGONAL BINNING
Density was calculated using hexagonal binning. In other words the number of PS points per hexagonal area was calculated. Hexagonal areas were used for this for the following reasons:
1. Hexagons have symmetry of nearest neighbours, which is lacking in square bins.
2. Hexagons are the maximum number of sides a polygon can have for a regular tessellation of the plane, so in terms of packing a hexagon is 13% more efficient for covering the plane than squares. This property translates into better sampling efficiency at least for elliptical shapes.
3. Lastly hexagons are visually less biased for displaying densities than other regular tessellations. For instance with squares our eyes are drawn to the horizontal and vertical lines of the grid.

The object oriented S plus statistical package was used to calculate the hexagonal density. In particular the ‘hexbin’ function was applied to the raw data, all settings were left as default, meaning the data would be binned into 30 hexagons in the east-west direction.

PLOTTING OF HEXAGONAL DENSITY DATA
The ‘hexbin’ function in S Plus creates a .bin file. In order to view this file in a meaningful manner the ‘plot’ function in S Plus was used to create greyscale graphs showing the density. This was done in two ways:
1. All eight teams were plotted using the same X and Y-axis values. This allowed a direct comparison on the density and spatial coverage of the teams datasets (Figure 22).
2. All eight teams data were plotted using different X and Y-axis values. This allowed the area of presentation to be maximised for each team (Figure 23).

DENSITY FOR DIFFERENT AREAS
Density was then examined more closely with respect to the land cover type. Firstly an overall figure for density (defined here as number of PS points per kilometre squared) was computed. In order to ensure comparability between the teams the same spatial area had to be examined for each team. However the spatial area processed by each team varies. It was therefore necessary to identify an area that all teams had processed.

IDENTIFICATION OF AREA FOR DENSITY STUDY
For this part of the study the pre-processed data provided by BRGM was used for teams 2-7. Data for teams 1 and 8 were from the most recent update as provided by ESA on the FTP server.
Counts of persistent scatterers in hexagonal cells - scaled for approx. 30 cells (east-west)

Figure 22: Overall density, each team displayed on the same X, Y scale.

Counts of persistent scatterers in hexagonal cells - scaled for approx. 30 cells (east-west)

Figure 23: Overall density, each team displayed with different X, Y scales.
The PS data for teams 2-7 received from BRGM were in a MapInfo format. These were converted to ESRI shapefiles using the Universal Translator tool in MapInfo. Data for teams 1 and 8 were acquired from the ESA FTP server as .csv tables. These were converted to a DBF IV format in Excel before opening in an ArcGIS 9.1 project. The ArcGIS 9.1 project had been previously set up with a Lambert III cartographic reference system. The orthorectified aerial photograph of the study area was also loaded into the project.

It was necessary to identify an area, which was covered by PS data from all teams and had the mixture of urban and rural areas we were looking for. An initial look at the spatial coverage of all 8 teams shows that teams 6 and 8 data cover the smallest area. Therefore the datasets from teams 6 and 8 were displayed as different coloured points (team 6 = yellow, team 8 = red) over the orthophoto, as in Figure 24.

MEASURING THE DENSITY

A rectangular area covering as much of the 3 datasets as possible was then digitised. Figure 25.

Figure 24: ArcGIS project with the data for team 6 (yellow points) and team 8(red points) overlaying the orthophoto.
The area of this rectangle was computed using the ‘Calculate Areas’ tool in the ‘Utilities’ section of the ‘Spatial Statistic Tools’ within ArcToolbox. This area was then entered into an MS Excel table, such as in Table 3 below, which had been set up to compute the density of areas when the relevant information was entered.

<table>
<thead>
<tr>
<th>Team</th>
<th>No. PS in rectangle</th>
<th>No. PS in urban area</th>
<th>Density of rectangle</th>
<th>Density Urban</th>
<th>Density Rural</th>
<th>Area rural (km)</th>
<th>Area of rectangle (km)</th>
<th>Area urban (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>481</td>
<td>481</td>
<td>14.20</td>
<td></td>
<td></td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3055</td>
<td>3055</td>
<td>90.20</td>
<td></td>
<td></td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6472</td>
<td>6472</td>
<td>191.08</td>
<td></td>
<td></td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6857</td>
<td>6857</td>
<td>202.45</td>
<td></td>
<td></td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4872</td>
<td>4872</td>
<td>143.84</td>
<td></td>
<td></td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3973</td>
<td>3973</td>
<td>117.30</td>
<td></td>
<td></td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10497</td>
<td>10497</td>
<td>309.92</td>
<td></td>
<td></td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>22638</td>
<td>22638</td>
<td>668.38</td>
<td></td>
<td></td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: A MS Excel table which was created to calculate the density for different regions and land cover types.

The next step is to find how many PS points form each team lie within this rectangular area. This can be done in ArcGIS using the selection tool, from the main menu, and selecting features from each team that are completely within the rectangle previously digitised. This selection is then saved to a new layer file. Once the attribute table is opened the number of points from that team, that fall within the rectangle can be seen, Figure 26.
Figure 26: The red points are the points selected as being in the rectangle, the tabular data shows how many points are present. These points were saved out to a new file for future use.

Once we know the number of points selected, this figure can be entered into the Excel. Table 4. Since we are interested in the variation of PS density between Urban and rural areas it is necessary to define these areas. The aerial photograph was used as a basis to digitise polygons that had an urban land cover type. These polygons were restricted to the rectangular area previously defined as in Figure 27.

<table>
<thead>
<tr>
<th>Team</th>
<th>No. PS in rectangle</th>
<th>No. PS in urban area</th>
<th>No. PS in Rural area</th>
<th>Density of rectangle</th>
<th>Density Urban</th>
<th>Density Rural</th>
<th>Area rural (km)</th>
<th>Area of rectangle (km)</th>
<th>Area urban (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>481</td>
<td>481</td>
<td></td>
<td>14.20</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3055</td>
<td>3055</td>
<td></td>
<td>90.20</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6472</td>
<td>6472</td>
<td></td>
<td>191.08</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6857</td>
<td>6857</td>
<td></td>
<td>202.45</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4872</td>
<td>4872</td>
<td></td>
<td>143.84</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3973</td>
<td>3973</td>
<td></td>
<td>117.30</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10497</td>
<td>10497</td>
<td></td>
<td>309.92</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>22638</td>
<td>22638</td>
<td></td>
<td>668.38</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>33.870164</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Excel Table with area of rectangle and number of points within the rectangle entered for each team. The density of points within the rectangular areas has been calculated for each team.
The area of the urban areas was computed using the ‘Calculate Areas’ tool in the ‘Utilities’ section of the ‘Spatial Statistic Tools’ within ArcToolbox. This area was then entered into a MS Excel table, such as in Table 5 below.

The same procedure as for the rectangular area was used to select the PS points from each team that fell within these urban polygons. The number of points in the urban areas was then added to the Excel table, as in Table 3. Once the values for the areas of the urban polygons and the number of PS points that fall within the polygons are entered a density for the urban areas is derived. Also subtracting the urban area from the rectangular area produces the rural area, similarly subtracting the number of PS points found within the urban area from the number within the rectangular area gives the number of points within the rural area. From these two derived numbers a density for the rural area can also be calculated.

<table>
<thead>
<tr>
<th>Team</th>
<th>No. PS in rectangle</th>
<th>No. PS in urban area</th>
<th>No. PS in Rural area</th>
<th>Density of rectangle</th>
<th>Density Urban</th>
<th>Density Rural</th>
<th>Area of rural (km)</th>
<th>Area of rectangle (km)</th>
<th>Area of urban (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>481.00</td>
<td>393.00</td>
<td>88.00</td>
<td>14.20</td>
<td>33.62</td>
<td>3.97</td>
<td>22.18</td>
<td>33.87</td>
<td>11.69</td>
</tr>
<tr>
<td>2.00</td>
<td>3055.00</td>
<td>2356.00</td>
<td>699.00</td>
<td>90.20</td>
<td>201.57</td>
<td>31.51</td>
<td>22.18</td>
<td>33.87</td>
<td>11.69</td>
</tr>
<tr>
<td>3.00</td>
<td>6472.00</td>
<td>5490.00</td>
<td>982.00</td>
<td>191.08</td>
<td>469.70</td>
<td>44.27</td>
<td>22.18</td>
<td>33.87</td>
<td>11.69</td>
</tr>
<tr>
<td>4.00</td>
<td>6857.00</td>
<td>5639.00</td>
<td>1218.00</td>
<td>202.45</td>
<td>482.45</td>
<td>54.91</td>
<td>22.18</td>
<td>33.87</td>
<td>11.69</td>
</tr>
<tr>
<td>5.00</td>
<td>4872.00</td>
<td>4241.00</td>
<td>631.00</td>
<td>143.84</td>
<td>362.84</td>
<td>28.45</td>
<td>22.18</td>
<td>33.87</td>
<td>11.69</td>
</tr>
<tr>
<td>6.00</td>
<td>3973.00</td>
<td>3430.00</td>
<td>543.00</td>
<td>117.30</td>
<td>293.46</td>
<td>24.48</td>
<td>22.18</td>
<td>33.87</td>
<td>11.69</td>
</tr>
<tr>
<td>7.00</td>
<td>10497.00</td>
<td>8829.00</td>
<td>1668.00</td>
<td>309.92</td>
<td>755.37</td>
<td>75.20</td>
<td>22.18</td>
<td>33.87</td>
<td>11.69</td>
</tr>
<tr>
<td>8.00</td>
<td>22638.00</td>
<td>13704.00</td>
<td>8934.00</td>
<td>668.38</td>
<td>1172.45</td>
<td>402.76</td>
<td>22.18</td>
<td>33.87</td>
<td>11.69</td>
</tr>
</tbody>
</table>

Table 5. All the values are now complete.
VISUAL ASSESSMENT OF THE DENSITY

To gain a visual appreciation of how land cover types affect the density of the PS points it was necessary to produce a georeferenced density grid for each team. This density grid could then be overlaid with the urban areas identified above.

To produce the density grids, the ‘Spatial Analyst’ extension in ArcGIS was used. One of the options in this extension is the ‘Density’ option. The following parameters are required:

1. Input data. This is the PS point dataset.
2. Population field. This is left at ‘None’ we could enter a different field from the table if we wanted to examine the density of different attributes. This is not relevant to this situation.
3. Density Type, can be set to ‘Kernel’ or ‘Simple’. Simple was used, as this is a measure of points per unit area whereas kernel is a weighted measure of points within a mathematically defined surface.
4. Search Radius, was set to 100, this is the distance to search for points from the centre of each cell in the output.
5. Area units. This was set to Square meters
6. Output cell size. This was set to 100 meters. 100-meter cells were chosen as these allowed a good compromise between detail of results and processing speed.

The result is a gridded dataset, with a value every 100m. Since we were instructed not to make comparisons between teams it was decided to display each teams data with a colour ramp applied a with a 2 standard deviation stretch, Figure 28. This display method highlighted which part of a team’s dataset had the highest density. These density grids were displayed with the urban areas polygons allowing a comparison to be made between the two datasets.

DENSITY COMPARISON WITH AREAS OF UNSTABLE GROUND

Areas of unstable ground were identified from the DiffSAR results produced by BRGM. This georeferenced dataset was used and a polygon for unstable ground digitised.

This polygon was then overlaid with the density results as shown in Figure 29. This allowed a visual assessment on the effect of ground instability on the density of PS points.
Figure 28: 100m$^2$ density results for Team 3 displayed with blue representing highest density and red lowest. Areas with no colour have no PS points present. White polygons are urban areas.

Figure 29: 100m$^2$ density results for Team 3 displayed with blue representing highest density and red lowest. Areas with no colour have no PS points present. White polygons are urban areas and the blue polygon is the area of ground motion as identified from the DiffSAR results.
A.8 Average geocoding errors

For each unique X_rY_r radar coordinate, the corresponding X_lY_l Lambert coordinates were extracted from the eight team databases. Between 1 and 8, XY Lambert coordinate pairs can be found for a single radar coordinate, because all teams did not use the same set of X_rY_r radar coordinates. Some teams use sub-pixel information, resulting in radar coordinates containing a decimal fraction. To match these coordinates to the integer coordinates of the other teams, they were rounded towards the nearest integer. The error which is introduced by this inevitable rounding is 0.7 times the size of a radar pixel at maximum.

For each set of two teams, the average distance between Lambert coordinates, the standard deviation and the number of pairs were recorded (see tables below).

Table 6: Average distance (in m).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>40.9</td>
<td>45.9</td>
<td>48.6</td>
<td>32.5</td>
<td>136.8</td>
<td>41.8</td>
<td>70.1</td>
</tr>
<tr>
<td>2</td>
<td>40.9</td>
<td>-</td>
<td>17.9</td>
<td>16.9</td>
<td>16.7</td>
<td>112</td>
<td>21</td>
<td>44.6</td>
</tr>
<tr>
<td>3</td>
<td>45.9</td>
<td>17.9</td>
<td>-</td>
<td>11.8</td>
<td>8.2</td>
<td>97.7</td>
<td>21.9</td>
<td>39.3</td>
</tr>
<tr>
<td>4</td>
<td>48.6</td>
<td>16.9</td>
<td>11.8</td>
<td>-</td>
<td>17.6</td>
<td>95.5</td>
<td>18.7</td>
<td>43.5</td>
</tr>
<tr>
<td>5</td>
<td>32.5</td>
<td>16.7</td>
<td>8.2</td>
<td>17.6</td>
<td>-</td>
<td>112.5</td>
<td>20.8</td>
<td>47.1</td>
</tr>
<tr>
<td>6</td>
<td>136.8</td>
<td>112</td>
<td>97.7</td>
<td>95.5</td>
<td>112.5</td>
<td>-</td>
<td>116</td>
<td>75.4</td>
</tr>
<tr>
<td>7</td>
<td>41.8</td>
<td>21</td>
<td>21.9</td>
<td>18.7</td>
<td>20.8</td>
<td>116</td>
<td>-</td>
<td>44.4</td>
</tr>
<tr>
<td>8</td>
<td>70.1</td>
<td>44.6</td>
<td>39.3</td>
<td>43.5</td>
<td>47.1</td>
<td>75.4</td>
<td>44.4</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7: Standard deviation of distance (in m).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>23.5</td>
<td>20</td>
<td>24.2</td>
<td>18.1</td>
<td>27</td>
<td>24.7</td>
<td>21.1</td>
</tr>
<tr>
<td>2</td>
<td>23.5</td>
<td>-</td>
<td>9.6</td>
<td>14.3</td>
<td>12.3</td>
<td>21.3</td>
<td>14.8</td>
<td>18.3</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>9.6</td>
<td>-</td>
<td>5</td>
<td>8.3</td>
<td>19.8</td>
<td>3.2</td>
<td>13.2</td>
</tr>
<tr>
<td>4</td>
<td>24.2</td>
<td>14.3</td>
<td>5</td>
<td>-</td>
<td>15.3</td>
<td>19.3</td>
<td>7.3</td>
<td>17.8</td>
</tr>
<tr>
<td>5</td>
<td>18.1</td>
<td>12.3</td>
<td>8.3</td>
<td>15.3</td>
<td>-</td>
<td>19.8</td>
<td>6.8</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>21.3</td>
<td>19.8</td>
<td>19.8</td>
<td>-</td>
<td>23.4</td>
<td>23</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>24.7</td>
<td>14.8</td>
<td>3.2</td>
<td>7.3</td>
<td>6.8</td>
<td>23.4</td>
<td>-</td>
<td>21.4</td>
</tr>
<tr>
<td>8</td>
<td>21.1</td>
<td>18.3</td>
<td>13.2</td>
<td>17.8</td>
<td>14</td>
<td>23</td>
<td>21.4</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 8: Number of pairs of radar-targets which have been compared per tie.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>84</td>
<td>368</td>
<td>405</td>
<td>208</td>
<td>109</td>
<td>608</td>
<td>219</td>
</tr>
<tr>
<td>2</td>
<td>84</td>
<td>-</td>
<td>12348</td>
<td>4075</td>
<td>1547</td>
<td>381</td>
<td>11432</td>
<td>2649</td>
</tr>
<tr>
<td>3</td>
<td>368</td>
<td>12348</td>
<td>-</td>
<td>14573</td>
<td>10009</td>
<td>1758</td>
<td>39293</td>
<td>6811</td>
</tr>
<tr>
<td>4</td>
<td>405</td>
<td>4075</td>
<td>14573</td>
<td>-</td>
<td>2985</td>
<td>1499</td>
<td>21793</td>
<td>3231</td>
</tr>
<tr>
<td>5</td>
<td>208</td>
<td>1547</td>
<td>10009</td>
<td>2985</td>
<td>-</td>
<td>1057</td>
<td>10972</td>
<td>3113</td>
</tr>
<tr>
<td>6</td>
<td>109</td>
<td>381</td>
<td>1758</td>
<td>1499</td>
<td>1057</td>
<td>-</td>
<td>2113</td>
<td>781</td>
</tr>
<tr>
<td>7</td>
<td>608</td>
<td>11432</td>
<td>39293</td>
<td>21793</td>
<td>10972</td>
<td>2113</td>
<td>-</td>
<td>7725</td>
</tr>
<tr>
<td>8</td>
<td>219</td>
<td>2649</td>
<td>6811</td>
<td>3231</td>
<td>3113</td>
<td>781</td>
<td>-</td>
<td>7725</td>
</tr>
</tbody>
</table>