EuroSDR-Project
Commission 1
“Sensors, Primary Data Acquisition and Georeferencing”

“Reliability of Direct Georeferencing:
An Overview of the Current Approaches and Possibilities”
Final Report on Phase 1

Report by Jan Skaloud
École Polytechnique Fédérale de Lausanne (EPFL)
# ABSTRACT

1 INTRODUCTION

1.1 Background

1.2 Motivation

1.3 Limits

1.4 Outline

2 RELIABILITY AND INTEGRITY

2.1 Definitions

2.1.1 Reliability

2.1.2 Integrity

2.2 Application to DG

3 DG IN GENERAL

3.1 The method

3.2 Technology suppliers

3.3 Overview

4 GNSS

4.1 Current situation

4.2 Available technologies

4.2.1 RAIM

4.2.2 SBAS

4.2.3 GPS signal modernization

4.2.4 GLONASS and Galileo

4.2.5 PPP

4.2.6 CP-DGPS

4.2.7 Network differential techniques

4.2.8 Local and nation-wide networks

4.2.9 Differential atmosphere

4.3 Summary

5 INERTIAL SENSORS AND ESTIMATION METHODS

5.1 Current situation

5.2 Available technologies

5.2.1 The enabling technology

5.2.2 Sensor life expectancy

5.2.3 Sensor redundancy

5.2.4 GPS/INS integration

5.2.5 FDE in Kalman filtering

5.2.6 FDE in Artificial Neural Networks

5.2.7 Limits of GPS/INS and complementary methods

5.3 Summary

6 INTEGRITY AND COMMUNICATION

2
Abstract

After some initial hesitations, the direct georeferencing (DG) of airborne sensors by GPS/INS is now a widely accepted approach in the airborne mapping industry. Implementing DG not only speeds up the mapping process and thus increases the productivity, but also opens the door to new monitoring applications. Although the system manufacturers tend to claim that DG is a well established technique and no longer a research topic, the technology users often encounter pitfalls due to undetected sensor behavior, varying data quality and consistency. One could almost claim that the reliability of DG is the Achilles’ heel of this otherwise revolutionary approach in civil airborne mapping. EuroSDR has recognized this problem and would like to address it in several phases. First phase of this effort are some preliminary investigations, charting the current situation and making suggestions for further research. The investigations are divided into the following technology fields: GNSS, inertial sensors and estimation methods, integrity and communication, calibration and integrated sensor orientation. Each field describes the current situation with respect to DG and discusses additional existing possibilities. These do not claim to be complete or exhaustive; however, they claim to address the essential features, methods and processes, the combination of which could increase the reliability of DG substantially without setting large side penalties.

1 Introduction

1.1 Background

Within the last decade, the application of Direct-Georeferencing (DG) has brought a small revolution into the mapping industry by driving down the cost of mapping products and speeding up the production cycle. At the same time, it has enabled the practical introduction of sensors such as lasers, line scanning cameras, and radar systems into civil airborne mapping. Although DG can now be considered as a well established industrial method, there remain a number of open questions related to its reliability and/or data quality control (QC). These concern both the clients and manufactures, as they are often related to instrument- or method redundancy which influences the cost of a system and the speed of the production.

1.2 Motivation

In 2005, EuroSDR initiated a preliminary investigation into the reliability of direct georeferencing that shall help the future institutional activities in this area. The initial project phase aims at understanding the current situation and sketching an overview of the used or available approaches and technologies related to this topic; it will serve as a base for further decisions. The institutional ambitions can be summarized by two points:

- No commission has ever made a mapping system.
- No mapping or quality standards have ever been made without a commission.
1.3 **Limits**

The study is limited in time and resources and, therefore, its primary aim is to be rather informative than exhaustive. Also, part of the 'situation map' was drawn using responses from technical suppliers to a questionnaire. Unfortunately, in some critical cases, no or very limited responses were given. (The author sincerely thanks to those who took the time and effort for replying!) This may eventually distort the given picture in some way, hopefully not decisively.

1.4 **Outline**

After giving some initial definitions, an overview of the current situation is presented. The individual parts of the long chain of DG information flow are treated separately. Each part starts with a problem identification that is followed by a summary of available technologies and an ‘estimate’ of currently used approaches.

# 2 Reliability and Integrity

## 2.1 Definitions

### 2.1.1 Reliability

Reliability has various interpretations. In the DG context it mainly refers to
- the controllability of observations, that is, the ability to detect blunders and to estimate the effects that undetected blunders may have on a solution;
- the probability of a system to function under stated conditions for a specified period of time.

The former is often decomposed into internal and external reliability. Internal reliability relates to the amount of gross error in an observation, not detectable at a certain probability level while the external reliability relates to the effect of non-detectable blunders on the estimated quantities (for example coordinates).

The latter context can be expressed mathematically as \( R(t) = \int_{-\infty}^{\infty} f(x)dx \) where \( f(x) \) represents the Failure Probability Density Function (FPDF) and usually refers to physical signal failures.

### 2.1.2 Integrity

By definition, integrity is a measure of trust which can be placed in the correctness of the information supplied by the total system. Integrity includes the ability of the system to provide timely warnings to

*the user when the system should not be used for the intended operation [1].* The integrity risk is the probability of an undetected (latent) failure. The systems of highest ambitions are of high reliability (i.e. never break down) and high integrity (i.e. a brake down is immediately detected) but in principle there can be systems of high integrity but low reliability or vice versa.
2.2 Application to DG

The controllability of observations is closely related to redundancy that significantly increases system reliability; it is often the only viable means of controlling. However, redundancy comes at higher price either due to additional components, signals, or processing methods. Augmenting reliability by redundant observations will be the main interest of this study that follows the individual sensors and data fusion.

The total failure of equipment or one of its parts is usually easily detectable while the occurrence of an unexpected error or performance degradation may be more difficult to notice. Such degradation may bias the DG solution outside its estimated accuracy. This aspect of reliability as investigated here is therefore understood more in a sense of the trustworthiness of the estimated performance.

3 DG in General

3.1 The method

Georeferencing can be defined as a process of obtaining knowledge about the origin of some event in space-time. Depending on the sensor type, this origin needs to be defined by a number of parameters such as time, position, attitude (orientation), and possibly also the velocity of the object of interest. When this information is attained directly by means of measurements from sensors aboard the vehicle, the term direct georeferencing is used. In other words, DG comprises a long process of information flow that involves acquisition, synchronization, processing, integration, and transformation of measurement data from navigation (GPS/INS) and remote sensing instruments such as frame or line scan cameras, lasers or radars. The term of DG is sometimes understood as a one-directional data flow from GPS/INS to the mapping sensor(s). When there is a common treatment or a feedback between remotely sensed data and navigation parameters, the term of Integrated Sensor Orientation (ISO) is used.

3.2 Technology suppliers

The limited field of options that existed only some years ago is diversifying rapidly. This may bring some advantages to the users in terms of pricing; however, it also increases the risk in terms of quality. As will become quickly evident, the purpose of this study is NOT to list or evaluate the technology suppliers! The investigations are limited to the conceptual level of available and future technologies (used/not used) and therefore no concrete references to providers are given.

3.3 Overview

Most of the technology suppliers have identified that the successful integration of DG into the mapping process requires knowledgeable users as much as good software functionality. Hence, at least the serious players periodically organize and encourage training courses. Those also provide well established workflows; however, these are often optimized for a particular system. The notion of the system- or process reliability is currently traded for a less clear definition of data quality control (QC). The QC comes at different stages of data processing, however, often with substantial delays that do not allow calling a mapping mission successful with good confidence at landing or at the end of the
day. The general lack of redundancy (and thus reliability) in navigation instrumentation (both at the physical and signal level) needs to be compensated by ISO. Hence the requirements on QC as well as on additional important issues such as system calibration are very different depending on whether ISO is used or not.

4 GNSS

4.1 Current situation

In most scenarios, the position of the airborne carrier is determined by one dual frequency GPS receiver (and one antenna) on board of a vehicle. The trajectory accuracy is usually improved off-line by carrier-phase differential data using forward/backward processing and ambiguity determination/validation for one or more base stations. In situations like platform stabilization, real-time GPS/INS integration is performed, however, not in the differential mode. This means that the final answer on data accuracy and reliability cannot be obtained with high confidence during the data acquisition phase. Moreover, possible occurrences of local signal distortions affecting both the GNSS code and phase measurements remain difficult to control and become apparent only later in ISO (bundle adjustment, LiDAR strip adjustment). In general, the reliability measures are replaced by “data QC” that is introduced on different levels. It comprises checks on grammatical (physical) and semantic (validity) aspects of the signal, the geometric situation in real time, and processing residuals in post-processing. Overall, the GNSS-derived position is the decisive factor for trajectory accuracy at lower frequencies (<0.1Hz). With all the progress in carrier-phase differential techniques its application usually marks the mission outcome (i.e. success or failure).

4.2 Available technologies

4.2.1 RAIM

In terms of physical reliability and integrity, there is a great difference between aviation-certified GPS receivers and the consumer GPS receivers [2]. Apart from the resistance to harsh environment, electromagnetic interference, clearly defined low-dB tracking scenarios and time to first-fix, the avionic receivers use standardized methods for Fault Detection and Exclusion (FDE). The whole process is also known as Receiver Autonomous Integrity Monitoring (RAIM). It requires a minimum of 5 satellites and uses the probability density function and minimum bias or worst bias with fixed or variable threshold [3]. It is based on the Bayesian approach of mixing probability density functions (nominal & failure case) and weighted by their probabilities of occurrence [4]. RAIM can provide alarm during the flight but it is useful only if the operator has access in real time to this information and the possibility to act in order to correct the problem; for example by collecting new data or by changing the trajectory. RAIM is not a standard option in consumer GNSS receiver technology [5] and it is not clear to which extent this is used in the acquisition phase of the DG process.

4.2.2 SBAS

The Satellite-Based Augmentation Systems (SBASs) currently comprise WAAS (Wide Area Augmentation System) covering good part of North America, EGNOS (European Geostationary Navigation Overlay Service) covering Europe and parts of its surroundings, and MSAS (Multi-
Transport Satellite Based Augmentation System) covering part of Asia and Pacific including the Japanese territory. The signal of these systems is interoperable and they offer satellite signal integrity monitoring in flight [1] as well as estimates on ‘normal’ deviations in GNSS signals (such as atmospheric delays, satellite clock-, and ephemeris errors). In other words, such a system ‘flags’ obviously erroneous measurements and computes quality metrics for the others that are broadcast along with the corrections. It is important to note that the decision what to do with this information is left upon the user receiver. Again, the receiver behaviour using SBAS-input is regulated only in case of avionic equipment [6]. The positioning accuracy using the suchlike augmented GPS signal is reported to be 1 to 2 meters vertically and around 1 meter horizontally for EGNOS [7], [8], [9] and slightly worse for WAAS [10] under optimal conditions. Although this accuracy is better than standalone GPS it is still insufficient for most DG applications. Nevertheless, the concept of monitoring the integrity and quality of the code-measurements can well contribute to the DG acquisition phase. Most likely, this has not yet been fully exploited for various reasons.

4.2.3 GPS signal modernization

The modernization of the GPS signal comes in different phases. First, L2C (C/A code on L2) is being introduced on the IIR-M block of satellite. Although one SV has been in orbit since September 2005, the nominal 24 satellites providing this signal are not scheduled before 2012. The main advantages of this enhancement are an improved interference resistance and tracking capability (~3dB higher). Some L2C-ready receivers are already available on the market. The impact on trajectory accuracy and thus DG performance is not expected to be significant before the introduction of the 3rd civil carrier frequency (L5) on the Block IIF and Block IIIA satellites. This will take even longer to materialize.

4.2.4 GLONASS and Galileo

The GLONASS constellation is currently enjoying a new boom (13 active SVs in 2005) that is scheduled to continue until reaching a complete constellation of 26 SV in 2012. Its impact on DG applications has been limited up to now but may gain importance once more SVs become available. The proposed signals for Galileo should bring benefits for code multipath mitigation thanks to ‘faster’ codes (steeper slopes of the correlation peaks) and data-free sidelobes. Since end of December 2005, the first experimental Galileo satellite has been transmitting its signal in space [11] that is currently under the process of validation. Its full constellation is scheduled for 2010; however, the ‘five years goal’ has been shifted already several times in the past. Hence, the improved reliability through redundancy of systems, satellites and signals is not expected to happen any earlier before 5-7 years from now.

4.2.5 PPP

The Precise Point Positioning (PPP) is a concept of GPS positioning using data from a single GPS receiver and precise satellite orbit and clock information generated by the International GPS Service (IGS). This technique is reported to achieve decimetre or sub-decimetre accuracy without the need for processing any GPS reference station data. PPP can make use of single- [12] or dual- [13], [14] frequency carrier-phase measurements. The drawback is usually a considerable delay in algorithm initialization and sometimes the method stability as well as the need for an on-line access to IGS-derived products that come also with a certain delay. Nevertheless, there is a significant potential for this already commercially available technique for DG applications with relaxed accuracy requirements or those executed over large remote areas. The saving comes in terms of simplified logistics. Some DG-related research projects focus on this methodology.
4.2.6 CP-DGPS

The double differencing (DD) of GPS carrier-phase (CP) and code data is the most common technique in trajectory estimation that allows achieving cm- to dm- level positioning accuracy under ‘normal’ conditions. For this end, the best estimate of the DD carrier-phase ambiguity needs to be computed (usually by the LAMBDA technique [15] or other least-squares methods) and validated [16]. Although the theory and practice of this process has progressed considerably, open questions still exist especially in the validation area [17]. The expected performance of ambiguity resolution is measured by its success rate given by the probability distribution of the integers. The results are different if the integers are computed based on geometry-free or geometry-based models. Consider an example in the case of DD and the geometry-based model supposing optimal tracking conditions and a short baseline: the instantaneous success rate is ~99.90% with 6 satellites used. However, local disturbances such as multipath, radio interference or ionospheric disturbances can quickly jeopardize this theoretical value. Another limit affecting the ambiguity fixing/reliability is the baseline length between the base station and the rover. Up to distances of 5 km, it is possible to work (at least theoretically) with L1 receivers. For <15 km baseline lengths, a L1/L2 data processing is necessary. For <30 km baseline lengths, additional data/products have to be added to the L1/L2 DD carrier-phase ambiguity fixing resolution [18]. This is usually achieved via a network of reference stations.

4.2.7 Network differential techniques

The network differential GPS techniques fall into one of three categories: (1) measurement domain, (2) position domain, and (3) state-space domain. Category (1) algorithms provide the user with corrections from a reference station or a weighted average of corrections from a network of reference stations. In approach (2), the user derives independent positions using corrections from separate reference stations. A weighted average of these solutions is then computed. The disadvantage of algorithms of group (1) and (2) is a degradation of accuracy with distance from the network centre. Moreover, (2) is not very well suited for ambiguity resolution although it is probably the most common approach used in DG applications (in post-processing). Its alternative is the true multi-baseline processing that is more common in studies of geodynamic phenomena. In this approach, all baselines are computed together, taking into account the inter-baseline correlations which arise from observing a GPS network simultaneously [19]. The approach (3) tries to estimate the real physical parameters as satellite clocks and orbits, reference station tropospheric- and clocks errors. However, its success depends not only on correct modelling but also on parameter observability and correlation. The ionospheric delays can also be modelled from dual-frequency reference station data for single-frequency end users. The recently adopted RTCM 3.0 standard foresees transmitting the reference measurements rather than the corrections or parameters to the user, who is finally left with the option to decide how to exploit them [20]. Hence, some previously investigated concepts of the trajectory reliability within the GPS network may become more practical to apply [21].

4.2.8 Local and nation-wide networks

Only a few GPS receivers offer RTK solutions that work with several bases simultaneously, i.e., the user can set up a mini-network without implementing servers and other network-specialized tools. In one particular case, the firmware of the receiver allows three modes. The first mode selects the best (nearest) base and works with it. The second default mode works with all (up to three) bases independently and provides a weighted solution. The third mode works with all three simultaneously inside the triangle provided the rover belongs to it (firmware-based instant Virtual Reference Station - VRS).
The nation-wide networks have most applications in terrestrial or maritime domains. Many European states are already covered by such systems in total of their territories. The provided correction rates of up to 1Hz are sufficient for the expected flight dynamics when using GPS/INS integration. Their main product are the real time and post-mission corrections, mostly provided as ‘nearest’ or ‘VRS’ modes [22]. Unfortunately, neither of these modes is well suited for trajectories that stretch over larger areas as the base needs to be frequently re-selected to prevent too long baseline lengths. Although some networks propose area-correction parameters (FKP), their derivation uses proprietary (and thus non-transparent) methods where reliability measures cannot be added without difficulties. Hence, the ground reference station measurements are usually applied off-line using the previously mentioned approach (2). The situation for DG applications can, however, improve when all reference data become available to the rover as proposed in the master-auxiliary messages concept [20]. The major challenge will then remain in establishing a robust and fast communication link between the network and the carrier.

4.2.9 Differential atmosphere

The differential atmosphere is obviously not a technology but rather a serious problem that is worth mentioning separately. Its situation is somewhat special as it can be solved through modelling with few parameters that are, however, rarely observed in practice. The avionic applications of DG involve important height differences between the airborne and reference GPS antennas that bias the trajectory in height when the delay due to the tropospheric refraction is not modelled correctly. If the actual temperature and pressure profiles differ from those assumed by the model (as is often the case), the magnitude of such biases is at least 5-10 cm per 500 m of height difference. Some models are better than the others, but most of the popular ones yield satisfactory results when fed with appropriate meteorological data. Although digital sensors of this type are cheap and available they are rarely exploited and almost never placed on the carrier! The research activities around atmospheric effects on GPS signals mainly focus on ionosphere modelling with parameters derived from monitoring networks. The tropospheric refraction is usually modelled as a combination of the tropospheric zenith delay and a mapping function. Recently, NOAA (National Oceanic and Atmospheric Administration, USA) started an experimental product that provides tropospheric delay estimates based on a nation-wide GPS network [23]. A first step in mitigating the tropospheric effects is the use of meteorological data at the reference station. Better estimation of model parameters implies the use of environmental data collected at all travelled altitudes. It is therefore advisable to implement a residual tropospheric delay estimation using meteorological data recorded in the aircraft during the flight (not only ‘en route’ but also through the climbing/landing phase) to minimize the systematic errors due to local tropospheric effects that cannot be predicted by global model variables [24].

4.3 Summary

Table 1 summarizes the available GNSS methods with respect to the reliability measures and their ‘estimated’ usage in DG. The robustness of GNSS positioning as a method will improve with the increasing number of satellites and signals made available, however, the technologies available today could be better explored.
Table 1: Reliability techniques in GNSS.

<table>
<thead>
<tr>
<th>Segment/Error</th>
<th>Mitigation in RT</th>
<th>Mitigation post mission</th>
<th>Situation in DG</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV functionality</td>
<td>SBAS</td>
<td>DGPS analyses</td>
<td>Rarely done in RT</td>
</tr>
<tr>
<td>Rover functionality</td>
<td>RAIM</td>
<td>Too late</td>
<td>RT–usually only geometry</td>
</tr>
<tr>
<td>Base functionality</td>
<td>RT-Network</td>
<td>Network</td>
<td>Sometimes, no RT</td>
</tr>
<tr>
<td>Atmospheric Delays</td>
<td>SBAS</td>
<td>PPP, DGPS, CP-DGPS</td>
<td>via CP-DGPS, rarely in RT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff. Troposphere</td>
<td>Sensors at carrier + base(s)</td>
<td>Parameters not observed</td>
<td></td>
</tr>
<tr>
<td>Multipath/Interference</td>
<td>Receiver and antenna hw/sw design</td>
<td>Follows the evolution</td>
<td></td>
</tr>
<tr>
<td>Long Base</td>
<td>Multi-base processing, Master-Auxiliary</td>
<td>Not optimal, no RT</td>
<td></td>
</tr>
<tr>
<td>Ambiguity</td>
<td>RTK</td>
<td>CP-DGPS</td>
<td>Separated per base, no RT</td>
</tr>
</tbody>
</table>

5 Inertial Sensors and Estimation Methods

5.1 Current situation

Although the use of inertial technology in life-critical navigation and guidance applications requires the employment of several (redundant) inertial measurement units (IMU), DG exploits (almost exclusively) only one sensor. Should the unit start malfunctioning, the technology providers rely on detecting obvious failures within the hardware (in real-time) and the detection of eventual performance degradation via the integration with GPS data and its post-mission analysis. The conventional GPS/INS integration tools usually cannot identify sensor degradation from incorrect stochastic/model assumptions without the interpretation of an experienced user. In other words, the models and estimation methods used in DG are generally well optimized for expected sensor behaviour but not for the marginal cases.

5.2 Available technologies

5.2.1 The enabling technology

In the context of DG, the primary role of the IMU is in the determination of orientation. The use of GPS/INS integration can be seen as a self-calibration technique for the gyros (the calibrated accelerometers are also needed for that) and a high-frequency interpolator of the GPS position. The inertial technology has been evolving for over fifty years. The most promising technologies enabling the direct measurement of the camera's orientation came with the concepts of ring laser gyros (RLG) and fiber optic gyros (FOG), as well as the later evolution of strapdown dry tuned gyros (DTG) and quartz rate sensors. In general, the sensors of each technology span several orders of magnitude in terms of precision. As a rule of thumb, their precision is proportional to sensor cost and size [25]. The general trend is to rather use smaller and cheaper sensors that rely on calibration by GPS data. The potential of orientation accuracies for today’s most popular sensors is summarized in Table 2. The majority of the numbers indicated in the table have been confirmed experimentally during numerous testing.
### Table 2: Inertial attitude determination performance with GPS aiding.

<table>
<thead>
<tr>
<th>Time</th>
<th>Navigation grade (usually RLG)</th>
<th>Tactical grade (usually FOG, DTG)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>roll, pitch (deg)</td>
<td>yaw (deg)</td>
</tr>
<tr>
<td>1 sec</td>
<td>0.0008 - 0.0014</td>
<td>0.0008 - 0.002</td>
</tr>
<tr>
<td>1-3 min</td>
<td>0.0014 - 0.003</td>
<td>0.004 - 0.005</td>
</tr>
<tr>
<td>longer time</td>
<td>same as over 1-3 min but manoeuvre-dependent</td>
<td></td>
</tr>
</tbody>
</table>

#### 5.2.2 Sensor life expectancy

The life expectancy of an IMU is usually characterized by its MTBF (Mean Time Between Failures). The users and also the system providers are sometimes less careful about the life-expectancy of the inertial components. At least one provider (and the unfortunate clients) was surprised when the incorporated IMU with the officially stated low MTBF values of 500 hours (originally conceived for missile guidance) manifested its nominal life cycle already within the first year of service in DG applications. Typically, the MTBF figures for IMUs applied in DG exceed 10000 hours. The total failure of some component (not a slow degradation) is usually detected by the IMU hardware and communicated via a predefined message, the interpretation of which, however, is left to the user.

#### 5.2.3 Sensor redundancy

A redundant IMU (internally, in terms of sensors) is composed of more than three accelerometers and three gyroscopes. One approach is to combine the inertial observations in the observation space to generate a ‘synthetic’ non-redundant IMU; a second approach is to modify the inertial mechanization equations to account for observational redundancy. The latter may have some economical benefits as it does not require ‘doubling’ of all sensors. On the other hand, doubling or tripling all critical components is most likely the simplest, but not necessary the most economic way for fault detection and isolation. Although the concept of sensor redundancy is a common way for increasing the system reliability in avionics [26, 27], this method is relatively novel in DG [28] and also not available in commercial systems.

#### 5.2.4 GPS/INS integration

The inertial system is integrated with GNSS because it cannot navigate accurately in stand-alone mode for extended periods of time due to the rapid accumulation of systematic sensor errors. Besides, an INS can successfully bridge the absence of GNSS signals (due to whatever reason) or smooth its short-term fluctuation. Nonetheless, the traditional GPS/INS integration cannot be considered as a good replacement of sensor redundancy and fault detection for the following reasons: First of all, GPS and inertial sensors do not sense the motion dynamics in the same spectral bands. Second, the integration is usually performed within a Kalman Filter (KF) that is often engineered to trust the inertial senor more than the GPS in case of unpredicted disagreement. In other words, the KF is configured to reject GPS measurements outside the predicted interval of confidence that is built upon the models. As these models are tuned for the expected stochastic behaviour of the sensors, they are not prepared to react correctly under unexpected conditions.
5.2.5  FDE in Kalman filtering

The chosen architecture of a KF influences not only the optimality of estimation but also the ability of Fault Detection and Exclusion (FDE). In principle, the KF can be of centralized, decentralized or federated architecture and with/without adaptive design. The centralized KF integrates the data from all available sensors on the measurement level in an optimal manner. However, the fault detection within this architecture is difficult to achieve, even with the use of another (i.e., third) redundant sensor [29, 30]. On the other hand, the decentralised and federated KF have better competences in FDE. These concepts can be described as sets of more than one KF organized into successive integration. A sensor or a subsystem is associated with a sub-KF, the output of which is re-integrated in the overall KF. In the federated design, each sub-KF is accompanied with an index that expresses the trust given to its results (by an internal controlling mechanism). In principle, fault detection can be achieved by comparing the outputs of the different sub-KF [31, 32].

Adaptive filters work on possible modifications of the stochastic assumptions or model parameters [33, 34]. A bank of KFs can be dedicated to run on different stochastic assumptions and models [26]. Although it can be very computational-intensive, the filter banks can provide the FDE via the analyses of innovation or estimate history even for tightly-coupled GPS/INS integration [35]. The available DG systems, though, are usually limited to conventional GPS/INS integration (tightly or loosely coupled) and do not offer specialized fault-detection algorithms.

5.2.6  FDE in Artificial Neural Networks

More recently, the theory of Artificial Neural Networks (ANN) has been applied to the navigation-system modelling and fault detection. The ANN concept is based on a training process by which a set of coefficients are determined, usually without a physical meaning. The disadvantage of this concept in GPS/INS integration is that different motion scenarios require different training procedures and any abrupt change in motion may trigger an alarm that can erroneously be considered as a fault [36]. Again, this technique is not known to be used in DG applications.

5.2.7  Limits of GPS/INS and complementary methods

There is no such thing as a perfect instrument and, despite its undoubted power, the integration cannot completely eliminate all possible errors. In other words, the data integration handled by a Kalman filter/smoother cancels only the non-overlapping part of the sensor's error budget, i.e. the observable errors. Thus the ‘band width’ of the error cancellation may overlap only partially with the motion of interest as a function of instrument type and precision and the dynamics of an aircraft. For that reason, de-noising inertial data prior to mechanization has proven in some cases to be indispensable for attitude determination and effective procedures have been developed for that purpose [37]. Another significant portion of the residual orientation errors is most likely to be affected by the quality of the in-flight alignment. Usually, the filter/smoother keeps on refining the attitude of the inertial platform all along the flight. The strength of this process is its ability to decorrelate the misalignment errors from other error sources and is enhanced when sufficient dynamics are encountered (strong correlation among the desired parameters lowers the trust or the reliability in the estimated performance measures). Its weakness remains in the susceptibility to be influenced by the changes of the accelerometer errors and unmodelled part of the gravity field. Both influences appear as wrongly sensed accelerations that are ‘eliminated’ by (numerically) re-adjusting the previously aligned platform. Dropping the coupling with the accelerometers is possible once the platform is aligned and high accuracy gyros are available (i.e. 0.002-0.01 deg/h). As the high frequency part of the anomalous gravity field is likely to remain unmodelled, this concept may be appealing for certain types of
applications when operating over a ‘rough, unknown’ gravity field or when flying along survey lines at constant velocity.

5.3 Summary

In general, the failures and malfunctioning in a GPS/INS solution can be detected and corrected for, or eliminated, by adopting one or more of three possible concepts: (1) sensor redundancy, (2) functional- and error-model modifications, (3) and application of advanced estimation methods. Although centralised KF have proven to provide better estimates, their fault detection capabilities are inferior to the decentralised and federated architectures. However, the centralized KF can be use for fault detection in a setup where a bank of filters of different stochastic assumptions is run in parallel and redundant sensors are provided. In principle, sensor redundancy is a necessity, i.e., without it only ‘massive errors’ or ‘stop-of-operation’ can be quickly detected. Although life-critical applications require triple redundancy as the minimum for the detection of failures and malfunctions, this may seem bit of luxury in DG domain. On the other hand, the evolution of inexpensive MEMS sensors may quickly remove such economical constrains. It also depends on whether it is sufficient to identify a faulty operation within a particular application, or whether exclusion and measurement replacement needs to be provided. In both cases, the currently available DG systems have little to offer as the (additional) sensor redundancy is practically non-existing and FDE not adopted.

6 Integrity and Communication

6.1 Current situation

As formerly defined, integrity asks for the alarm in real-time or with a predefined latency. The bulk of DG applications require the fusion of data collected on the carrier and on the ground (e.g. by CP-DGPS). The prerequisite of integrity-factor calculation on all levels is therefore the establishment of reliable (intra-system) communication links between all important components. This approach is generally applied in avionics by expensive and redundant infrastructure while it is almost non-existing in DG. As the demand on trajectory accuracy in DG applications is usually higher, the approaches pursued in avionics can only be regarded as complementary. On the other hand, the time latency is less critical in DG and therefore the publicly available methods of mobile communication represent an interesting solution.

6.2 Available technologies

6.2.1 The problem of distributions

From the theoretical and practical point of view, the verification theories applied in integrity monitoring require the use of Gaussian distributions. However, most of the error sources in GNSS (and inertial sensors) do not follow a Gaussian distribution. Worse, some error sources are not always zero mean, especially when observed over a short period of time. The navigation community addresses such problems by ‘overbounding’ [1]. Extension of this concept to the whole complexity of DG is far from being trivial.
6.2.2 Avionic approach

Today, only the integrity of code measurements can be estimated efficiently. SBAS, GBAS (Ground-Based Augmentation System) and ABAS (Aircraft-Based Augmentation System) are used in the computation of the integrity level. These techniques include or can be complemented with RAIM and GPS/INS integration. Unfortunately, the applications of DG require a higher level of accuracy than provided by code measurements. Nevertheless, some conceptual approaches or existing integrity algorithms can most likely be applied to carrier-phase data and to GPS/INS integration.

6.2.3 Pseudolites

The integrity concept exploiting CP-DGPS technology has been proposed for the CAT-III landing with the help of ground beacons – pseudolites (pseudo-satellites) [38]. The application-based limits when broadcasting integrity messages were identified as multipath and radio interference [39, 40]. The concept of pseudolites is also better suited for locally-limited applications and thus not for DG in general.

6.2.4 TCAR

The integrity verification of phase measurements in real time requires redundancy in the computation of the positioning solution. Ideally, a second (redundant and independent) solution is computed. An approach could be based on the new civil signals of GPS and Galileo and the TCAR- (Triple- or Three-) Carrier-Phase Ambiguity Resolution) [41, 42] or FAMCAR techniques (Factorized Multi-Carrier Ambiguity Resolution) [43]. Thus, over-determination could be provided by a multi-carrier solution and a “traditional” CP-DGPS solution with the possible help of GPS/INS integration.

6.2.5 Communication technology

Communication links are required for the real-time transmission of GPS corrections or measurements and integrity information. The transmission of this information ranges from (geostationary) satellites (SBAS) to terrestrial wireless data transmission techniques. For CP-DGPS, radio, cellular terrestrial, satellite, and wireless transmission are compared in Table 3 based on the availability of the communication network, the provided bandwidth, the range, and the cost of the communication link. The integrity requirement in avionics asks for a priority communication link, which is perhaps not necessary in DG. Furthermore, the communication link must not interfere with the GNSS signals (this issue is critical for satellite communication [44]).

<table>
<thead>
<tr>
<th></th>
<th>Radio</th>
<th>GSM</th>
<th>GPRS/UMTS</th>
<th>SatCom</th>
<th>802.x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proprietary</td>
<td>-</td>
<td>+</td>
<td>+/−</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Data rate</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+/−</td>
<td>+/−</td>
</tr>
<tr>
<td>Availability</td>
<td>+/−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Range</td>
<td>+/−</td>
<td>+</td>
<td>+</td>
<td>±</td>
<td>+/−</td>
</tr>
<tr>
<td>Multi-channel</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Cost</td>
<td>+/−</td>
<td>−</td>
<td>−</td>
<td>±</td>
<td>+/−</td>
</tr>
</tbody>
</table>

Table 3: Comparison of communication links.

Radio transmission is used for the traditional RTK applications. Its inconvenience for DG applications is the low range due to the low transmission power. As (physical) weight (essentially for power
supply) is not critical here, the range can be increased using higher transmission power as long as the legal requirements are fulfilled. GSM proved to be limited by its data rate of only 9.6 kbps that corresponds approximately to 5 Hz of dual-frequency measurements from one reference station [45]. The network setup or the arrival of new civil GNSS signals further increases the demand on data throughput. The availability of GSM (as well as GPRS and especially UMTS) decreases in rural regions of European countries and these technologies are not ‘generally’ available in many countries. The problems related to cell registration and hand-over are known to occur for fast moving carriers, such as aircrafts. GPRS has higher data bandwidth as compared to GSM. Unfortunately, the unexpectedly reduced and varying data throughput have proved to be an important inconvenience for kinematic CP-DGPS applications [46]. The newly implemented UMTS technology can handle even higher data transfer rates; however, the transmission is usually handled by ‘bursts’ of packets and therefore has varying latency.

The principle advantage of satellite communication based on Low Earth Orbiting- (LEO) satellites (the availability of GEOS is highly reduced in mountainous regions) is their availability. Some systems are limited to 9.6 kbps (for Globalstar), while the broadband service providers (e.g. Skybridge, Teledesic) offer somewhat higher data rates. The 802.x wireless communications techniques (e.g. 802.11x, Bluetooth, ZigBee) are of very short range with the exception of a directive array.

### 6.3 Summary

A complete integrity concept for DG would need to face a challenging communication problem when operating over large areas or remote regions. Although the use of dedicated infrastructure would be technically feasible, it is more realistic to foresee sub-optimal or hybrid systems that make a better use of the available technologies such as SBAS, nation-wide GPS networks, and existing communication systems. In smaller projects, the use of radio transmission seems (still) to be the most appropriate communication means for passing GNSS data or corrections and – perhaps in the future – integrity messages.

### 7 Calibration and Integrated Sensor Orientation

#### 7.1 Current situation

In the context of reliability, the Integrated Sensor Orientation (ISO) currently represents the security net for the DG. The net casting can be wider or narrower according to the sensor-type, accuracy requirements, and performance of navigation data. Moreover, the use of ISO is inevitable for the system calibration. The calibration process is not standardized and each technology provider offers some tools for this purpose. The comprehension of the technology’s principles and limits, the ‘savoir faire’, and the judicious data handling seem to be more important than the functionality of a particular tool. Therefore, the users are sincerely invited to follow a specialized formation either in academia or with system providers.

Although the use of ISO requires additional work compared to DG, the process of image orientation is no longer ‘doubled’ in practice (e.g. derivation of exterior orientation with and without GPS/INS). Instead, the complementarities of methods are put upfront as in the self-calibrating GPS/INS-AT where a fast and almost automated transfer of homologous points can be achieved. Although the methods of integrated adjustments have room for improvements, this space is much larger for LiDAR or SAR than for the frame- or line-based sensors.
7.2 Available technologies

7.2.1 System calibration in general

What is understood by system calibration is the process of finding the relations in position (lever-arm), orientation (boresight), and time (synchronization) between the sensors. The calibration of systematic effects in the imaging/ranging sensors (e.g. parameters of camera interior orientation, LiDAR range-finder offset) can be made either separately or within the same process. The concepts of state-space estimation (KF in GPS/INS) and bundle adjustments (AT) have the ability to accommodate and estimate additional calibration parameters. However, doing so may cause severe correlation among the variables and hamper the reliability of the whole process. Hence, independent methods and parameter separation is recommended whenever feasible.

7.2.2 Lever-arm calibration

The lever-arm calibration is a typical example of the previous note on parameter calibration where ISO is not indispensable but (often) used. The lever-arm effects can be correctly modeled and thus calibrated within the KF and/or within the bundle adjustment. However, even good observation conditions cannot match the accuracy of determination by independent geodetic (tachometric) means. Even worse, the lever-arm parameters are often strongly correlated with other systematic errors, e.g., of the inertial or the GPS observations [47]. Nevertheless, the software-driven approach of adding additional parameters represents often the most economic and convenient way for the user that is unaware of the related dangers.

7.2.3 Boresight calibration for frame and line-based sensors

Contrary to the lever-arm, the calibration of the boresight requires performing an ISO for attaining sufficient accuracy. The related problems have been addressed by many investigations [48-52]. The situation for frame-cameras is relatively well understood, although some conceptual approaches are better than the others and possibilities for improvements exists [47]. Conceptually, the situation is not very different for line-based scanners when ‘pushbroom’ image blocks are formed and adjusted [53].

7.2.4 Boresight calibration for LiDAR

Contrary to well-developed approaches to boresight estimation, the correct recovery of the LiDAR-IMU misalignment is considerably more complicated. The adopted approaches are usually based either on physical boundaries or cross-sections [54, 55], DTM/DSM gradients [56] or signalized target points [57]. These procedures, while functional, are recognized as being sub-optimal since they are labor-intensive (i.e., they require manual procedures), non-rigorous, or they provide no statistical quality assurance measures. The more rigorous class of calibration procedures or strip adjustments uses the modeling of systematic errors directly in the measurement domain [58, 59], yielding practical and adequate results with good de-correlation between all parameters [60].

7.2.5 Synchronization

The synchronization between the sensors in airborne applications should be performed with a maximum time tolerance of 0.1ms. Previously, varying time delays used to be a problem especially
when existing image sensors were retrofitted with DG equipment; however, these problems are hopefully eliminated by proper electronic design in the era of new digital instruments. Control of timing can, e.g., be performed by imagery overlaps flown from opposite directions.

7.2.6 Sensor interior orientation

The calibration procedures for digital sensors were recently very well documented by the corresponding EuroSDR-initiated activity [61]. The situation remains less clear for LiDAR [62] and almost proprietary in case of airborne SAR.

7.2.7 Transformation of EO to national coordinates

The choice of a mapping frame and projection is often an underestimated factor causing tensions or distortions. The non-Cartesian character of national (often conformal) projections is causing distortions when DG is performed without special modifications of the bundle adjustment software [63]. Until recently, the problem alleviation by modified transformation of GPS/INS-derived EO was not correctly addressed openly. Apart from the curvature of the earth, the main problem is that the basic equations of photogrammetry rely on a Cartesian reference frame. National mapping frames, however, are not Cartesian due to the length distortion encountered when projecting an ellipsoid into the plane [63, 64]. Further, national maps are often based on local geodetic datums that differ from the reference frame in which the GPS/INS solutions are obtained. There are, in principal, three different ways to solve these difficulties: (1) the photogrammetric restitution in a suitable tangential frame and the subsequent transformation of the complete scene to national coordinates, (2) the computation of artificial ground-control points and restitution based on their transformation (imitation of indirect georeferencing), and (3) the restitution directly in national coordinates. The latter approach requires special attention when coping with the earth curvature and the length distortion of the national map projection. A detailed investigation on all these aspects is found in [65].

7.3 Summary

The concept of ISO is very powerful in the reliability control and needed for system calibration. The main problems of this approach are: (1) the additional work that cannot be fully automated and therefore delays the delivery; and (2) the fact that it comes as a last step and therefore almost too late (from an economical point of view) if the decision to re-fly needs to be taken. The procedures for system calibration can be still improved and the best available methods are not always followed. The latter applies also to the use of DG in map projections and local coordinate systems. Open problems still exist especially in the context of calibration of LiDAR and SAR sensors.

8 Concluding Remarks

As the GPS/INS technology starts to represent the sole means of sensor orientation (DG) in many projects, the factors concerning its reliability are gaining importance. The reliability is closely related to sensor redundancy and system complexity and thus the overall system cost. However, the higher
‘upfront’ expenses for more reliable systems could be saved later when dropping current (and sometimes less reliable) methods of quality control, consistency checks, or the laborious process of integrated sensor orientation. This is even more evident if integrity concepts (related to reliability checks in real-time) can get introduced.

The chain of data flow in DG is long and the method is only as strong as its weakest link. In the context of reliability, this continues to be the carrier-phase differential GPS, especially over longer baselines. ‘Waiting for Godot’ (represented by Galileo in the context of the famous tragicomedy of two acts) is not most likely the approach to be taken as there is a number of possible technologies existing today, the combination of which may well alleviate the problem. Similarly, there are many possibilities for improvements within the GPS/INS integration itself, both on the hardware- and software level. Finally, although the sensor-to-sensor correlation/calibration problem is no longer a nightmare, the rigorous or standardized approaches are still far from common practice.

9 Acknowledgement

The author would like sincerely thank the doctoral students at EPLF-TOPO, namely Fadi-Atef Bayoud, Valérie Renaudin and Adrian Waegli for their important contributions in literature review and problem synthesis. The time and remarks of Dr. Klaus Legat are greatly appreciated when reviewing this report. Véronique Chazal is thanked for her assistance with final formatting and linking.
References


http://topo.epfl.ch/documents/EuroSDR/hatch00.pdf


http://topo.epfl.ch/documents/EuroSDR/burrell03.pdf


http://topo.epfl.ch/documents/EuroSDR/lehmann05.pdf

http://topo.epfl.ch/documents/EuroSDR/skaloud03.pdf


22


References by alphabetic order


24


Jacobsen, K., 2002. Transformation and computation of orientation data in different coordinate systems. OEEPE Official Publication, 43


25


Ober, P.B., 2000. Integrity according to Bayes. In: D.U.o. Technology (Editor), IEEE.
http://topo.epfl.ch/documents/EuroSDR/ober00.pdf


Optech, 2004. ALTM 30/70/100 user manual, Toronto, Canada.

Pervan, B.S. et al., 1995. High integrity GPS-based precision landing using beacon pseudolites, ISPA, Braunschweig, Germany, pp. 7.


http://topo.epfl.ch/documents/EuroSDR/skaloud05.pdf

http://topo.epfl.ch/documents/EuroSDR/skaloud03.pdf


http://topo.epfl.ch/documents/EuroSDR/talaya00.pdf

http://topo.epfl.ch/documents/EuroSDR/tempelmann05.pdf


