Rovira Más, F.; Banerjee, R. (2012). GPS data conditioning for enhancing reliability of automated off-road vehicles. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering. 227(4):521-535. doi:10.1177/0954407012454976.

GPS data conditioning for enhancing reliability of automated off-road vehicles

Francisco Rovira-Más[1] *, Ratul Banerjee[2]

[1] *Departamento de Ingeniería Rural y Agroalimentaria, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain*
[2] *Department of Electronics & Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, West Bengal 721302, India*

**Corresponding author*: Francisco Rovira-Más; Tel: +34 963877291; Fax: +34 963877299; email: frovira@dmta.upv.es

Abstract

The practical implementation of precision agriculture at a large scale has not occurred yet due to several reasons. Among them, the lack of uniformity and reliability in global positioning has discouraged many producers to adopt advanced solutions which, while considered to add a significant value to their production systems, cannot be incorporated before warranting minimum levels of long-term consistency. Although substantial improvements are constantly being introduced by receiver manufacturers, positioning errors can appear at the last stages of the localization process, resulting in inaccuracies and anomalies normally undetected by embedded quality filters. This article proposes an actuation protocol to enhance the robustness of GPS information for practical agricultural applications. The algorithm embodying this strategy merges partially-acquired raw strings into complete NMEA messages whose information fields are checked for consistency. Once data qualifies as stable, other logic filters are applied to reinforce the likelihood of obtaining proper locations. Extensive field tests demonstrated that the algorithm was able to discard most of erroneous positions due to typical GPS errors and poor signal reception in complex agricultural environments. However, the phenomena of coordinate quantization and random outliers were still present, which indicates that further redundancy is necessary to avoid unreliable outcomes. In this regard, supplementary consistency for GPS-based vehicle heading and speed anticipated positive results.

Keywords

GPS Error; NMEA Message; Precision Agriculture; Global Positioning; GPS Reliability; HDOP

1.- Introduction

Precision Agriculture (PA), a concept meant to revolutionize the root and basis of agricultural systems towards top efficiency and full sustainability, has reached a crossroad with most of the necessary technology available but with environmental and economic benefits yet unproven [1]. Ten years have passed since this conclusion was drawn, and the *status quo* is still the same; in a study to assess the potential use of precision agriculture (also denominated precision farming —PF—), Pölling et al. [2] confirmed that the implementation of PF technologies in practical farming has slowed in recent years on global scale compared to the initial euphoria of the mid and late 1990s. Although PA technologies have been experimented throughout the world [1], only a relatively small portion of farmers have practiced any type of PA-based technology. According to Bramley [3], the rate of adoption by growers of many crops remains low and, in some industries, is negligible. Given that the multiple benefits of PA are generally accepted, it is important to determine the real cause of this low rate of adoption. A perceived lack of benefit may have discouraged many potential adopters, but Bramley [3] concludes that due to the fact that the majority of agricultural producers are not "profit maximizers", there are other factors affecting their strategic decisions. The technical difficulties of efficiently handling novel technologies such as the Global Positioning System (GPS), Geographic Information systems (GIS), remote sensing, automatic control, and computing, all by managers instructed to manage the land rather than electronics and cybernetics, has been identified as one of the major factors, if not the most important, explaining low implementation rates. Because most experimental PA systems are map-based [1], global positioning is central to the practical implementation of precision agriculture. The understanding and proper manipulation of GPS information is essential for the efficient application of PA in the field. The real implementation of a system meant to work season after season, and typically not operated by an engineer, demands high consistency and resistance to failure. GPS —and other similar systems such as GLONASS, Galileo, or Compass— does not seem to have reached that status yet, at least for the case of free signal systems. Bramley [3] points out that the approximate $2500 annual subscription fee for optimized differential correction has been dissuasive for many Australian growers, even in high value crops like wine grapes. Therefore, improving the performance of affordable GPS systems could incentivize the generalized adoption of Precision Agriculture applications. However, even high accuracy systems such as RTK-GPS are not completely free of hindrances; not surprisingly, Mitchell [4] claims there exists a long-time desire for a universal and robust means to refer global coordinates for extended agricultural use, as winning technologies in the future will depend on the farmer who will sift through what is new and choose only what works. This paper intends to contribute to such demand by getting into the core of GPS messaging and delivering better positioning estimates.

In spite of the difficulties mentioned above, precision agriculture and other GPS-based applications are in expansion. A study based on recent patents on off-road automatic navigation [5] showed that GPS-based technologies became predominant in the 2001-2008 period, probably motivated by the cancellation of selective availability in 2000 by the US Department of Defense. In agreement with this trend, Humphreys et al. [6] speculate that a growing segment of military and civilian global navigation satellite systems (GNSS) users will demand greater accuracy and reliability from their receivers than can be offered by single-frequency GPS, as the overwhelming majority of GNSS receivers fail when signals are blocked or jammed. The solution, however, is neither simple nor immediate, and absolute unchanging reference positions are not a realistic expectation [7]. Yet, not all precision systems always require the highest degree of accuracy, although it is necessary to maintain errors within pre-established limits, whatever span these limits may have, so that uncontrollable outcomes do not happen. Pölling et al. [2] report very high potential for western and central parts of Europe, and Zhang et al. [1] indicate the same trend for China. The traceability of chemical applications is becoming critical for wine producers, who need to keep an accurate account of sprayed rows after the appearance of disease or insect activity. This information is accessible only when geographical references are available with a GPS. Larzelere and Landers [8] found that, in such applications, the selection of GPS receiver/antennae is critical because accuracy is key. In fact, inexpensive solutions led to unacceptable results, but higher levels of sophistication were useful for row and positional location even with free signals. In similar fashion, the three-dimensional mapping system of agricultural scenes developed by Rovira-Más [9] resulted very sensitive to GPS coordinate misplacement, and the slight offset of a small number of points was enough to degrade the entire map. In addition to mapping applications, GPS receivers are also essential for automated navigation, where the absence of positioning data or their lack of precision may result in dangerous situations. This condition has been counterweighted with the combination of GPS and inertial measurement units (IMU), such as the position-velocity-attitude model [10] based on the Kalman filter to provide accurate positioning with a low-cost GPS combined with solid-state inertial sensors. With a similar purpose, Mizushima et al. [11] developed a low-cost inertial unit composed of three gyroscopes, two inclinometers, and a differential GPS that was capable of correcting the drift accumulated by the gyroscopes. The expectation derived from these examples showing the multiple benefits of global positioning provides a favorable environment to pursue the improvement of GPS reliability in automated off-road vehicles by directly actuating on the acquisition and morphology of GPS messages.

2.- Errors and problems with GPS in practical field applications

GPS errors can be analyzed from the communication technology standpoint or, alternatively, from a practical, user-oriented, point of view. The former comprises satellite ephemeris, clock, atmospheric, multipath, and receiver errors. The latter have been enclosed in the terms bias and drift errors [12]. In general, field robotics applications highly benefit from differential signals — usually from a DGPS receiver— because satellite ephemeris and clock errors are totally cancelled; however, the cancellation of atmospheric errors (ionospheric and tropospheric) degrades with distance, and both multipath and receiver errors cannot be mitigated by differential corrections [13]. The presence of these errors in particular will manifest to the user in the field as bias and drift errors. A favorable location of the receiver antenna in conjunction with premium components and optimized software results helpful to palliate the negative effect of multipath and receiver errors, but the complete suppression of these errors can never be assured with absolute certainty. Nevertheless, from the user standpoint, finding out whether the origin of inaccuracies is atmospheric or caused by unavoidable reflections is not essential; what matters is the early detection of positioning errors and, when possible, their correction. In fact, it is not feasible for the average user to decouple atmospheric effects, multipath reflections, and receiver noise. Rather, it seems more practical to classify errors according to the user perception, which in reality can correspond to a sudden and unexpected jump called *bias* as it represents a deviation from the expected, or a steady loss of positioning accuracy over time denominated *drift*. Therefore, drift errors are caused by continuous signal degradation whereas bias errors result in random misplacements of global coordinates. The appearance of both errors is not mutually exclusive, and for that reason, right after a random jump (bias) in the trajectory of a vehicle, coordinates may start to drift or vice versa. Typical GPS receivers used in agricultural applications may begin to drift after 15 minutes of continued use.

In general, there is no convenient way to anticipate multipath deviations or receiver noise, which in practice means that end users of this technology cannot predict when a bias jump is imminent or drift is about to start. Yet, certain information related to the quality of the signal being acquired can be known with every positioning message. In particular, two parameters facilitated by the GPS Control Segment are helpful to evaluate the quality of the signal: the number of satellites in solution and the dilution of precision (DOP). In theory, three satellites suffice to determine the location of a receiver, but uncertainties in the measurement of time require a fourth satellite, and in practice only solutions obtained from the triangulation of five or more satellites offer enough guaranties for field operations. As a result, it is necessary to know the number of satellites locked by the receiver for each calculated point. As a matter of fact, not only the number of satellites but their spatial distribution above the receiver is critical for getting precise estimates. The *dilution of precision* is a set of parameters that quantify how favorably satellites are distributed in the sky in relation to a specific user location. They can refer to time errors introduced by the clock (TDOP), position errors in the horizontal (East-North) plane (HDOP), and errors in altitude (VDOP). The majority of agricultural operations take place in flat terrains, and therefore the *horizontal dilution of precision* (HDOP) is typically the key parameter to track. Some other times, however, sloping terrains recommend the use of the *vertical dilution of precision* (VDOP) as well, which tends to be higher than the HDOP. As a rule, the lower the HDOP the better distribution of satellites above the GPS antenna. Common values for the HDOP range between 1 and 3.

With quality parameters —number of satellites and HDOP— associated to each position registered in the field, reliability might be taken for granted and the problem of bias or drift easily solved by just discarding position estimates whose number of satellites or HDOP do not reach an established threshold. This reasoning might work for some especial situations or during particular periods where conditions are ideal, but it cannot be considered a generalized solution. The map plotted in Fig. 1, for instance, represents the trajectory followed by a tractor along four rows of a vineyard. The path traced by the vehicle in the map properly indicates the actual course travelled in the field. However, out of the 218 points recorded, there exist 13 outliers that result in unrealistic jumps (bias errors), some of which over 20 meters. Although it only represents a 6 % of the points, this result is unacceptable for many real applications such as implement control, automatic navigation, mapping, or site-specific variable rate dosing. The special case of autosteering poses serious risks to the control of vehicles, which cannot afford offset errors of that magnitude. Nevertheless, the eradication of outliers would set the map ready to be used, which apparently can be done by identifying points with a low number of satellites and high HDOP rates. Surprisingly, the graph depicting the number of satellites (Fig. 2a) for all the points represented in Fig. 1 shows that there were never less than 8 satellites in solution, and similarly, the history of HDOP along the rows (Fig. 2b) was either 1.3 or 1.5. This outcome evidences that the deterministic assessment of GPS stability is not trivial, and given that many agricultural applications that rely on GPS positioning cannot progress unless an acceptable performance of GPS is assured in regular practice, the challenge of enhancing the robustness of GPS data in real time provides a strong motivation for conducting this research.

Fig 1. Outliers in a typical map of a vineyard showing the trajectory followed by a vehicle.

Fig 2. Number of satellites (a) and HDOP (b) for the test shown in the map of Fig. 1.

3.- Philosophy and algorithm for conditioning GPS messages

The enhancement of GPS reliability can be performed at different stages of the message transmission process, and a big effort is currently being made by researchers and receiver manufacturers to improve performance at the electronics and signal processing level [6]. However, while these improvements are always favorable to the achievement of precise operations in the field, they are not sufficient; any anomalous datum introduced at the last steps —including coordinate transformations— might ruin the usefulness of a map or the accuracy of an execution command. Therefore, consistency in the data must be maintained until the very end of the spatial referencing procedure. However, system designers must focus on the optimum use of positioning sensors rather than improving their rated performance, which falls in the scope of manufacturers and sensor developers. Once the GPS antenna has been installed in a favorable location of the vehicle and the receiver power turned on, the sequence of strings output by the positioning device via serial port constitute the raw material from which global coordinates have to be calculated. Fortunately, these strings are not issued in a free style, and the popular standard NMEA 0183 has been universally adopted by GPS practitioners. Created by the US National Marine Electronics Association (NMEA), this standard specifies several structures for GPS messages in ASCII code (simple text), where information about time, position, and signal precision are usually included. The particular fields composing each message depend on the specific NMEA identifier used. Agricultural applications mostly use GGA (Time, position, and fix) and VTG (Course and speed) messages. This research was conducted using the GGA message type, but the procedure developed and the results found are also valid for any other type of NMEA messages. Fig. 3 provides the specifications for GGA messages, and further details on other NMEA strings are available in Rovira-Más et al. [14]. An actual example of a GGA string can be examined in expression (1).

Fig 3. Components of GGA messages (Standard NMEA 0183).

"$GPGGA","112633.0","3928.995234","N","00020.233222","W","2","10","1.0","3.911" ,"M","52.122","M","10.0","207*6D" (1)

Before processing the information acquired with the GPS, it is essential to fully understand each single field of the message structure of Fig. 3, as not all of them are equally important for the applications considered here. All the NMEA messages start with a six-character identifier, formed with a dollar sign ($) opening the message and followed by two letters identifying the sender (GP for GPS) plus three letters labeling the type of message (GGA, VTG, etc). After the final character of the last data field there is an asterisk immediately followed by a two-digit checksum expressed as a hex number. Out of the 15 fields comprising GPGGA messages, the algorithm proposed requires the GPS time, the three geodesic coordinates (latitude, longitude, and altitude), the number of satellites in solution, and the HDOP. Note that data fields are separated by commas, and when information on a field is not available, two adjacent commas will appear in the message indicating a null field. This possibility allows the incompleteness of messages, which will also be subjected to other hindrances caused by serial port transfer jamming and computer processor sharing. At present, there are two frequencies available for GPS devices: 1 Hz and 5 Hz. As a result, it is not possible to increase the flow of incoming GPS data above 5 Hz. All these difficulties complicate the structure of the filtering algorithm, represented as a block diagram in Fig. 4.The following points summarize the essential highlights of the algorithm:

1. Extensive fieldwork shows that GPS frequencies of 1 Hz and 5 Hz are very low compared to other sensors, and the time used by the computer to transfer messages through the RS-232 port is often insufficient to assure the acquisition of the entire string. The algorithm features a *piece-by-piece data acquisition system* which reinforces string *piece-wise composition* and reduces the loss of primary GPS information.
2. The flow of incoming data is severely influenced by satellite constellation layouts, and many times it takes too long to lock an acceptable number of satellites. Given that practical applications cannot be completely deterred waiting for GPS messages, the algorithm introduces a number of *timing checks* at several levels in such a way that a reasonable flow for the main process is assured (time out allowance). GPS receivers are usually one of many sensors providing key information and should never monopolize the pace of the main control loop.

3. Because NMEA GGA messages are standard and unambiguously defined, there is an *a priori* knowledge of the expected strings, as shown in Fig. 3. This knowledge provides a means to check for *string consistency*; there must be 15 sections (or fields) in each complete message, and the expected length of each section can also be known beforehand. Thus, for instance, the latitude is the third field of (1) and Fig. 3, and it consists of four characters followed by the decimal point and a number of decimals; the first two numbers are the degrees (39) and the rest —until the next coma— are the minutes (28.995234). The longitude, likewise, is contained in the fifth field, but this time the degrees are represented by the first three numbers (000) and the minutes by the rest (20.233222). Following this rationale, corrupted strings can be detected by keeping track of the number of sections and their corresponding sizes for any string of characters delimited by two dollar signs marking the initial points of two consecutive messages. The 15 fields of a string do not need to be acquired all at once, but only strings with the proper structure and dimensions are further processed; the rest are discarded. The lengths of these fields are not universal and usually vary with the specifications of the receivers so that low-cost receivers typically provide less precision. Consequently, the string consistency routine must adapt to different types of receivers.
4. The strict filter applied to the format of the strings is able to remove wild outliers pointing at impossible locations, but there is still no warranty of getting correct localization data. For that reason, a further step focuses on the particular analysis of some fields for which domain logical boundaries can be established, that is, the implementation of *section-based logic domains*. Typical section-based thresholds were established for reasonable altitudes (altitudes between sea level and 1000 m), minimum number of satellites (usually 5), maximum HDOP allowance, realistic forward velocity (below 40 km/h for tractors), and so forth. This philosophy has been further developed in the 2012 version of the algorithm, in which VTG NMEA messages are processed simultaneously with GGA messages. As two speeds are determined by different sources, differentials of 30% in velocity are being used as indicators of potential errors.

Fig 4. Algorithm proposed for conditioning GPS messages.

Many GPS-based applications developed in the past have considered the own receiver quality indices to be enough to determine the acceptability of incoming localization data. This research intends to get deeper inside the transmission and morphology of GPS messages so that spurious data and any lack of consistency can be detected early enough to avoid unreliable measurements, with special emphasis on applications based on emergent technologies such as precision farming and field robotics. Although primary data coming from GPS satellites may be enhanced with high accuracy systems as proprietary differential signals (DGPS) or real-time-kinematic solutions (RTK-GPS), the majority of commercial systems used by average producers cannot afford such levels of accuracy, and as a result, the more efficient and reliable GPS positioning becomes, the better. Additionally, even sophisticated positioning systems are also susceptible of transferring corrupted strings whenever satellite constellations are adverse, reflections or signal blockage results critical, or the processing routine collapses.

4.- Equipment and field tests

The standard midsize tractor of Figure 5a was the vehicle chosen to test the algorithm and record data. It is equipped with a StarFire iTC^{TM} SF2 (Deere & Co, Moline, IL, USA) differential positioning system. Although SF2 correction signals with static accuracy of 25 cm and pass to pass accuracy of ± 10 cm are available under license, tests were performed using exclusively the free signal SF1, with a static accuracy of 75 cm and pass to pass accuracy of ± 33 cm. The algorithm developed to filter GPS messages was implemented in a mini computer onboard the tractor (AOpen MP945-D, San Jose, CA, USA), with a 1.8 GHz Core 2 Duo processor and 512 MB memory. A customized program written in C++ applied a coded version of the routine schematized in Fig. 4, easing the acquisition process with the graphic interface of Fig. 5b where the origin of coordinates and the output text files can be selected by the user. The control screen of Fig. 5b allows the visualization of the instantaneous position of the vehicle in the local tangent plane (LTP) system of coordinates, which requires the real time transformation from geodetic to LTP coordinates. The local tangent plane coordinates are *east*, *north*, and *altitude*, and besides of being more intuitive and easy to represent than geodetic coordinates, they permit the user to set a local origin of coordinates, avoiding oversized numbers that complicate the construction of information maps. Fig. 6 represents the position of a point P in both geodetic (a) and local tangent plane (b) coordinates. Although the real time transformation between both coordinate systems is critical for the successful implementation of the proposed solution, its step by step execution falls outside the scope of this article. It can be consulted in Rovira-Más et al. [14].

Fig 5. Testing vehicle (a), and control interface of the software developed (b).

Validation experiments took place in two diverse scenarios: a barren field in the vicinity of an urban area at sea level, where reflections and signal interferences were expectable; and a winery vineyard located in a rural environment about 650 m above sea level (Fig. 5a), where open fields provided a favorable environment for long testing periods. Multiple runs were driven under different conditions of traveling speed, vehicle orientation, time of day, atmospheric situation, and satellite constellation. The experiments were conducted in 2010 between May and August (Table 1).

Fig 6. Geodetic (a), and local tangent plane (b) systems of coordinates.

5.- Results & discussion

The first condition requested to the GPS data-conditioning algorithm was the capacity to identify and discard those locations with a low probability of being correct according to NMEA quality indices embedded in GPGGA messages. These indices are the number of satellites in solution (NOS) and the horizontal dilution of precision (HDOP), and were always tracked in this first-stage check. The intricacies of the proposed algorithm, apparent in Fig. 4, were designed to only deal with complete —and *a priory* healthy— strings from which meaningful values of NOS and HDOP could be extracted. Fig. 7a shows the trajectory followed by a tractor from the testing field to a workshop located inside a university campus. Along the approximate 500 m that separate the initial and destination points, there were large trees and a number of buildings which reflected and blocked satellite signals. A threshold of 5 or more satellites and HDOP inferior to 3 left several portions of the trajectory without information; yet it is always preferable to obtain an incomplete map rather than a wrong one. Fig 7b provides the NOS along the entire run and Fig 7c plots the HDOP over the same period. Notice the inverse profile between NOS and HDOP, and the minimum requirements met by all the points acquired by the tractor. The valid NOS ranged from 5 to 9 and the HDOP oscillated between 1.2 and 3. Only failing in one of the two restrictions was enough to reject the GPS message being processed. The application of the filtering routine resulted in an incomplete but sufficiently precise map.

Fig 7. Message filtration based on minimum NOS and maximum HDOP values: (a) trajectory; (b) NOS; and (c) HDOP.

According to the philosophy proposed, the first operation for conditioning the incoming data takes place at the message reception level, where primary strings are properly composed and checked for consistency. Then, quality indices NOS and HDOP establish a second level of filtration to remove unreliable messages. However, the important question of how precise and trustable will be the data that finally gets transferred to a map is still remaining. This query is not trivial, as positioning information that passes the previous filters is assumed to be correct, and unfortunately this will not always be the case. The run of Fig. 8 provides a real example of this problem. The trajectory of Fig. 8a (NOS = 8-9; HDOP = 1.4-1.6) was recorded moving north-east at 1.5 km/h, whereas Fig. 8b (NOS = 7-8; HDOP = 1.2-1.4) was traced by a tractor heading south-west at 5.5 km/h. Almost one month elapsed between both tests, and the testing sites are separated by 70 km in distance and 650 m in altitude. Yet, they show a common unexpected feature; motion along the north-south direction occurs at *discrete intervals* due to an *east-west drift* that *clusters* points horizontally. This phenomenon evidences that positioning points in the local tangent plane system of coordinates are being *quantized* in a *quanta* of points of up to 10 m in length (Fig. 8b). The consequences of this outcome are significant for the development of GPS-based vehicle applications; the calculation of vehicle heading and forward velocity from GPS coordinates, for instance, will lead to severe errors and potential dangerous situations. Furthermore, a magnified look at these results might reveal even worse consequences as the phenomenon of *cluster swapping*. Fig. 9 represents the trajectory traced by a tractor moving south-west at 4 km/h, with a permanent rate of 8 satellites in solution and an HDOP of 1.3 along the run. The quanta effect over clustered points along the east-west direction is noticeable for the entire run; however, the circles drawn in the plot highlight the occurrence of *cluster swappin*g, which states that for any three consecutive clusters at different levels, the one in the middle cannot be located either higher north or lower north than its bounding clusters. The consequences for heading calculations are obvious; two consecutive headings, one pointing north and the following pointing south, create a source of instability for determining vehicle states.

Fig 8. Quantization of coordinate points along the east-west direction.

Fig 9. Cluster swapping in local tangent plane coordinates.

Regardless of the filters and conditioning routines setup in cascade through which positioning messages are driven, there is no practical way to grant the absence of isolated random readings. Electromagnetic interferences or random data generated in the computer may originate meaningless values of key parameters. After a recording session held on June 29, 2010, in which the number of satellites had been high (over 7) and the HDOP low, the close inspection of data revealed that a few estimates had *randomly mutated* latitude 39 by latitude 29, which is, by all means, impossible. Just one single wrong number changed the position of the vehicle from Spain to Africa. Fortunately, this kind of errors happened very seldom, but they cannot be accepted for a true application even if they happen once. Mapping algorithms must include strong consistency filters before issuing definite information maps, prescription protocols, or site-specific automatic instructions. It is very valuable to be site-specific, but being wrong-site specific may induce serious problems. The path traced by the set of points of Fig. 10 reveals two outliers at points 13 and 26, in spite of a NOS between 6 and 9 (Fig 11a) and an HDOP value oscillating between 1.2 and 2.5 (Fig. 11b). As expected, both outliers appeared before point 30 when conditions of NOS and HDOP were less favorable according to Fig. 11; but on the other hand, point 26 was acquired with 8 satellites in solution and 1.3 HDOP, which are considered to be positive values for GPS localization. The negative effect of these two outliers can be appreciated in Fig. 12, which represents an estimation of the vehicle heading. The actual values of the heading were 80º going north-east, and –100º going south-west in the return pass. The plot shows how consistency strongly deteriorates due to the two outliers, resulting in 50º jumps that can never be acceptable regardless of the application pursued. The effect of these outliers on the calculation of the forward velocity were even worse, leading to jumps from the actual velocity of 5 km/h to unrealistic values of almost 25 km/h. Nevertheless, the occurrence of undetected outliers was very low when the total amount of tests is taken into account. Table 1 provides an overview of the experimental design and the results found. Notice that the presence of undetected outliers was induced by high demands of computer power when simultaneously running the GPS filtering routines and machine vision applications in the same computer. The maximum number of points represented in Table 1 is the direct application of a 5 Hz sampling rate to the duration of the tests Δt. This ideal sequence will never be reached due to the multiple processes that take place in the onboard computer, which not only acquires the NMEA strings from the RS-232 but also process them, displays the results, and records the important information in text files. As expected, the difference between the maximum possible number of points and the actual valid points registered increased as parallel processes were activated in the main processing loop. The repetitions were carried out over the same scenarios but the particular course traced by the tractor depended on each particular test, which ranged from short runs of about one minute to multiple-row paths of over 40 minutes.

Table 1. Experimental design and summary of results.

Test scenario	Repetitions	Test duration Δt (min)	Maximum nº points for Δt	Valid points	Undetected outliers
University roads	3	3.3; 4.3; 6.3	990; 1290; 1897	104; 658; 644	0; 0; 0
University field (1-15 June)	5	6.7; 5; 3.8; 16.3; 7.7	2008; 1508; 1146; 4881; 2309	1029; 773; 646; 2458; 1130	0; 0; 0; 0; 0
University field (16-30 June)	6	3.7; 1.8; 7.3; 1.8; 4.8; 1.9	1113; 549; 2180; 543; 1453; 568	612; 298; 1164; 294; 717; 292	0; 1; 0; 0; 0; 0
Vineyard (July, GPS mapping)	10	25; 8.4; 40.5; 46.2; 7.1; 3.4; 2.3; 3.4; 4; 7.4	7496; 2530; 12145; 13854; 2127; 1012; 690; 1032; 1185; 2208	2905; 1230; 5300; 1754; 405; 484; 397; 503; 622; 1095	0; 0; 0; 0; 0; 0; 0; 0; 0; 0
Vineyard (August, GPS + vision)	7	3.5; 1.2; 3.8; 1.3; 7; 3.7; 9.4	1055; 375; 1140; 385; 2085; 1120; 2810	67; 46; 81; 35; 137; 119; 218	2; 0; 2; 0; 11; 2; 13
Vineyard (Manual trigger vision)	4	10; 18.8; 1.6; 11	3000; 5640; 490; 3310	403; 915; 82; 571	7; 10; 0; 7

Fig 10. Random outliers found in the course traced by an agricultural tractor.

Fig 11. Quality indicators NOS and HDOP for the trajectory plotted in Fig 10.

Fig 12. Heading estimation for the trajectory represented in Fig. 10.

6.- Conclusions and lessons learned

The algorithm proposed in Fig. 4 to improve the quality of GPS messages based on the NMEA code was able to remove spurious data by assembling proper strings from consecutive acquisition loops. In addition, it incorporates several consistency checks to enhance reliability based on the number of satellites in solution, HDOP, altitude, heading, and traveling velocity. All these measures resulted in the practical strengthening of a global positioning system for agricultural applications, which may range from the basic registration of a vehicle’s course to the combination of the GPS receiver with more sophisticated algorithms for navigation or crop mapping and monitoring. However, although basic performance was achieved in general terms, experimental data revealed interesting facts for further investigation. First, while NOS and HDOP can be considered essential parameters to discard messages of low reliability, they cannot be taken as definitive evaluators to grant high reliability, as random mutations may appear over any stage of the positioning process. Second, the phenomenon of clustering along the east-west direction in discrete steps is important for working velocities below 5 km/h; therefore, according to each application developed, its effect may or may not impact the final results. And third, the appearance of random outliers, such as those shown in Figs. 1 and 10, need special treatment. A detailed exploration of the 13 outliers plotted in Fig. 1 revealed that the errors only affected the north coordinate transformation. However, since the same algorithm was applied to the entire series, and 94 % of the points were correctly determined, it is evident that the problem is not in the algorithm itself, but in the overload of computer processes that made the algorithm create artifacts at the time these 13 points were transformed. Furthermore, Table 1 shows that the algorithm was 100 % effective when the computer was not engaged in parallel high-cost computations, and only a few outliers appeared in severe demands of processing power. In any case, regardless of the origin of the fault, the main concern is the detection of conflictive points prior to their final delivery and utilization. GPS-based heading and velocity might bring vital information to assist in the real time detection of random outliers. Fig. 13 represents the estimated heading of the vehicle that acquired the information shown in Fig. 1. Forward runs had a heading of 80º, whereas return runs had a heading of –100º. The in-field estimated heading of Fig. 13 indicates stable values at the expected angles of 80 and –100, delimited by big changes in heading over the headland turns, which can be anticipated after taking 180º U-turns to change rows. What cannot be justified, however, is the set of isolated peaks for the heading recorded within the 100-meter straight rows, where vehicle orientation is approximately constant. These sharp peaks —of 50º or more— coincide with random outliers that point at unrealistic disruptions in the course followed by the vehicle; by keeping track of these peaks, the filtering system can be aware of potential outliers while map construction is still in progress.

Fig 13. GPS-based heading estimation for the trajectory plotted in Fig 1.

The embedded quality indices NOS and HDOP are essential to implement the data conditioning strategies outlined in this work. Although they represent different concepts, namely the number of satellites in solution and how homogeneously these satellites are placed with respect to each other, there exists a practical relationship that couples them, so that a low NOS implies a high HDOP and vice versa. The extent and mode of this relationship may be significant for consistency purposes. Obviously, a number of satellites cannot be associated with a unique HDOP because there are many ways in which satellites can spread out above the antenna of receivers, but some general boundaries can be defined according to practical experience. Fig. 14 identifies an area of high likelihood for the pair NOS-HDOP. This area is limited by two quadratic curves, whose mathematical equations are given by Eq. 2 and 3, and which are the result of a regression analysis conducted over 22 field experiments performed from June 2 to August 23, 2010. According to Fig. 14, given a number of satellites, a normal range for the HDOP can be predicted by drawing a vertical line at the current NOS and calculating the intersection with both bounding curves. This prediction can also serve as a resource to detect anomalies in applications featuring GPS-based localization.

$$[HDOP]_{min} = 0.0411 \cdot NOS^2 - 0.8046 \cdot NOS + 4.8714\ ;\ \forall NOS < 11 \quad R^2 = 0.9884 \qquad (2)$$

$$[HDOP]_{max} = 0.0786 \cdot NOS^2 - 1.5443 \cdot NOS + 8.80\ ;\ \forall NOS < 11 \quad R^2 = 0.9838 \qquad (3)$$

Fig 14. NOS-HDOP practical relationship.

The methodology described along this article was further tested and challenged in 2011 with an alternative low-cost GPS receiver (Garmin 18x 5 Hz, Garmin International Inc, Olathe, KS, USA). The reported phenomena of drift and coordinate quantization were also observed for the alternative receiver. Furthermore, jumps were present in both directions E-N and outliers appeared with more severity due to the lower capacity of the low-cost receiver. These outcomes corroborated the findings of the tests conducted in 2010 and proved that the problems then detected were not receiver-specific; however, the algorithm was equally efficient in the conditioning process of primary NMEA strings. The use of a less sophisticated receiver induced the improvement of the timing checks and waiting loops, as more control by the user on data input delays had to be implemented in order to let GPS strings enough time to be acquired through the serial port RS-232. Computer performance was found to be critical in the generation of errors; when parallel processes related to machine vision were highly demanding, the occurrence of outliers incremented. Consequently, the positioning engine of the vehicle must be considered and analyzed in conjunction with all the other tasks commanded to the processing computer, as the advantageous design of multitasking loops is decisive for the proper behavior of automated vehicles.

7.- Acknowledgements

The authors would like to thank Juan José Peña Suárez and Montano Pérez Teruel for their assistance in the preparation of the prototype vehicle, Verónica Sáiz Rubio for her help during most of the field experiments, and Luis Gil-Orozco Esteve for granting permission to perform tests in the vineyards of winery Finca Ardal. Our gratitude is also extended to the Spanish Ministry of Science and Innovation for funding this research through project AGL2009-11731.

8.- References

[1] Zhang, N., Wang, M. and Wang, N. Precision agriculture –a worldwide overview. *Computers and Electronics in Agriculture*, 2002, 36, 113-132.

[2] Pölling, B., Herold, L., Volgmann, A. and Wurbs, A., Werner, A. Assessing the potential use of Precision Farming-technologies in the EU. In Proceedings of International Conference on Agricultural Engineering AgEng, Clermont-Ferrand, France, 2010.

[3] Bramley, R. G. V. Lessons from nearly 20 years of Precision Agriculture research, development, and adoption as a guide to its appropriate application. *Crop & Pasture Science*, 2009, 60, 197-217.

[4] Mitchell, C. Down home and high tech. *Resource*, 2011, Special issue January/February, 17.

[5] Rovira-Más, F. Recent innovations in off-road intelligent vehicles: in-field automatic navigation, *Recent Patents on Mechanical Engineering*, 2009, 2, 169-178.

[6] Humphreys, T. E., Bhatti, J. A. and Ledvina, B. M. The GPS Assimilator, *Inside GNSS*, 2010, 5, 50-58.

[7] Strohman, R. and Mastin, T. Datums for a dynamic earth, ASABE Paper Nº 095776, 2009, American Society of Agricultural and Biological Engineers, St. Joseph, MI, USA.

[8] Larzelere, W. M. and Landers, A. J. Development of a spray monitoring system for a vineyard canopy sprayer. ASABE Paper Nº 1009012, 2010, American Society of Agricultural and Biological Engineers, St. Joseph, MI, USA.

[9] Rovira-Más, F. Global 3D Terrain Maps for Agricultural Applications. In Asim Bhatti (Ed.), *Advances in Theory and Applications of Stereo Vision*, 2011, InTech, 227-242, Available from: http://www.intechopen.com/articles/show/title/global-3d-terrain-maps-for-agricultural-applications.

[10] Guo, L., Zhang, Q. and Feng, L. A low-cost integrated positioning system of GPS and inertial sensors for autonomous agricultural vehicles. ASABE Paper Nº 033112, 2003, American Society of Agricultural and Biological Engineers, St. Joseph, MI, USA.

[11] Mizushima, A., Noguchi, N. and Ishii, K. Automatic navigation of agricultural vehicle using a low-cost attitude sensor. In Proceedings of Automation Technology for Off-road Equipment Conference: 98-106, Kyoto, Japan, 2004.

[12] Rovira-Más, F., Han, S., Wei, J. and Reid, J. F. Fuzzy logic model for sensor fusion of machine vision and GPS in autonomous navigation. ASABE Paper Nº 051156, 2005, American Society of Agricultural and Biological Engineers, St. Joseph, MI, USA.

[13] Grewal, M. S., Weill, L. R. and Andrews, A. P. *Global positioning systems, inertial navigation, and integration*, 2001, Wiley, New York.

[14] Rovira-Más, F., Zhang, Q. and Hansen, A. C. *Mechatronics and intelligent systems for off-road vehicles*, first ed., 2010, Springer, London.[Chapters 3 and 6]