Detection of Offsets in GPS Experiment (DOGEx)

Matt King[1], Simon Williams[2]

[1] School of Civil Engineering and Geosciences, Newcastle University • e-mail m.a.king@ncl.ac.uk
[2] National Oceanography Centre, Liverpool • e-mail sdwil@noc.ac.uk

Introduction

The accuracy of Global Positioning System (GPS) time series is degraded by the presence of offsets. If these are not detected and adjusted correctly they bias velocities, and hence geophysical estimates, and degrade the terrestrial reference frame. They also alter apparent time series noise characteristics as undetected offsets resemble a random walk process. As such, offsets are now a substantial problem in geodesy. A number of automated offset detection algorithms have been developed across a range of fields, and some of these are now being tested in geodetic time series

Here, we announce a community experiment in detecting offsets in GPS time series (DOGEx) and demonstrate the performance of a few of the offset detection approaches submitted to us so far.

Offset Categories		Offset is Present	Offset is Absent
Test or Metadata Shows	Positive	True Positive (TP) : Estimated in Least Squares	False Positive (FP) : Adds Additional Noise
	Negative	False Negative (FN) : Adds Additional Noise	True Negative (TN) : No Problems

Offsets submitted in a solution are sorted into 2 categories : True Positives (TP), if the offset is sufficiently close to a known offset; and False Positives (FP), if the offset does not correspond to a known offset. Actual offsets in the simulated dataset that have no TP associated with them are then classified as False Negatives (FN). An example if given in Figure 2 below left. The solution statistics are then based on the ratio of these parameters and the bias in velocity with respect to the "truth" using the true model. A perfect solution would have 100% TP, 0% FP, 0% FN and a 0 mm/yr velocity bias.

Methods

We have produced simulated 3-d GPS coordinate time series for 50 "sites". The simulated series contain realistic (and perfectly known) GPS signal, noise, offset frequencies and data gaps (e.g., **Figure 1**). Noise characteristics are modelled on that present in state-of-the-art GPS reprocessing solutions using a "white plus flicker" noise model, although the noise is not necessarily time-constant at each site

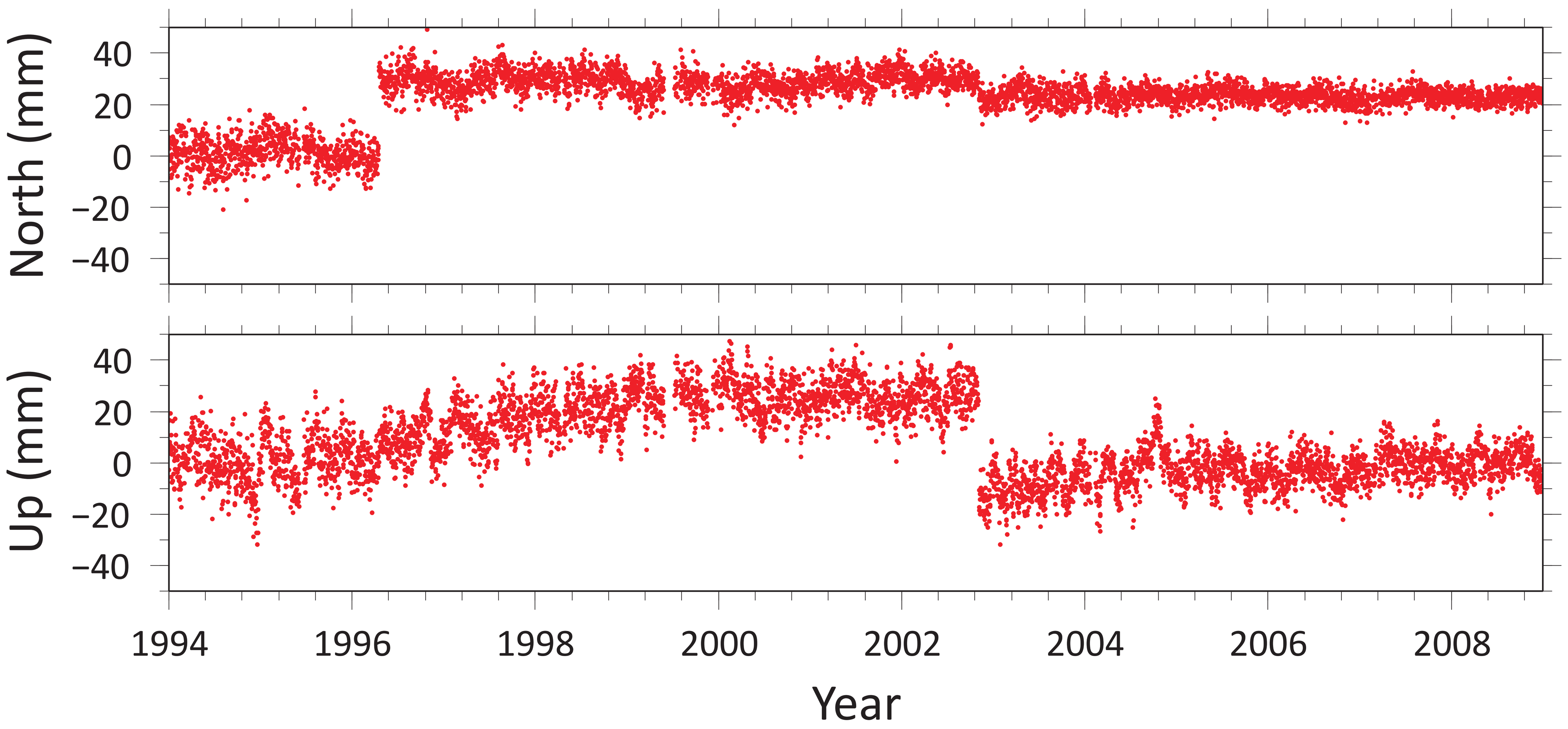


Figure 1 Example North and Up components of a synthetic DOGEx time series

Experiment Details

The DOGEx time series may be downloaded from

http://www.cost-es0701.gcparks.com/working-groups/working-group-3

The true offset times and site velocities will *not* be provided to the community

At regular intervals (IUGG, EGU, AGU) we will update the community on the best approaches and the effects of undetected or mis-detected offsets on GPS time series, velocities and apparent noise

Submitted Solutions

We now have a total of 22 solutions (15 not counting re-submissions after tweaking of the algorithms).

Solution	Version	Remarks	Thanks to
AIUBCOD1	1	Using the FODITS software	Luca Ostini and Rolf Dach
AIUBCOD2	2		
AIUBCOD3	3		
JPL_STP1	1	JPL Solution	Angie Moore, Susan Owen, Danan Dong, Sharon Kedar and Frank Webb
MAK1KF99	1	Shifting means hidden Markov model approach (Kehagias and Fortin, 2006; **K&F**)	Matt King
MAK2KF99	2		
MAK1PIEE	1	A maximum likelihood approach (Picard et al, 2005; **PIC**) with the number of segments chosen using one of two penalising functions (Lavielle, 2005; **LAV**, and Lebarbier, 2005; **LEB**) each considering **HOM**oscedastic and **HET**eroscedastic time series. PIEE = PIC_HET_LEB PIEA = PIC_HET_LAV PIOA = PIC_HOM_LAV PIOE = PIC_HOM_LEB	
MAK2PIEE	2		
MAK1PIEA	1		
MAK2PIEA	2		
MAK1PIOA	1		
MAK2PIOA	2		
MAK1PIOE	1		
MAK2PIOE	2		
MAK2CS3D	2	Iterative CUMSUM	
SDPWMANL	1	Offsets picked manually	Simon Williams
EJP_MANL	1	Offsets picked manually	Liz Petrie
MR_PCV_1	1	Automated solution	Marco Roggero
NOCLMANL	1	Offsets picked manually	Mark Tamisiea
ULGLFD01	1	Automated solution based on first difference time series	Norman Teferle, German Olivares Pulido
ULGLM001	1	Offsets picked manually	
BOSM_MLE	1	Offsets picked manually	Machiel Bos

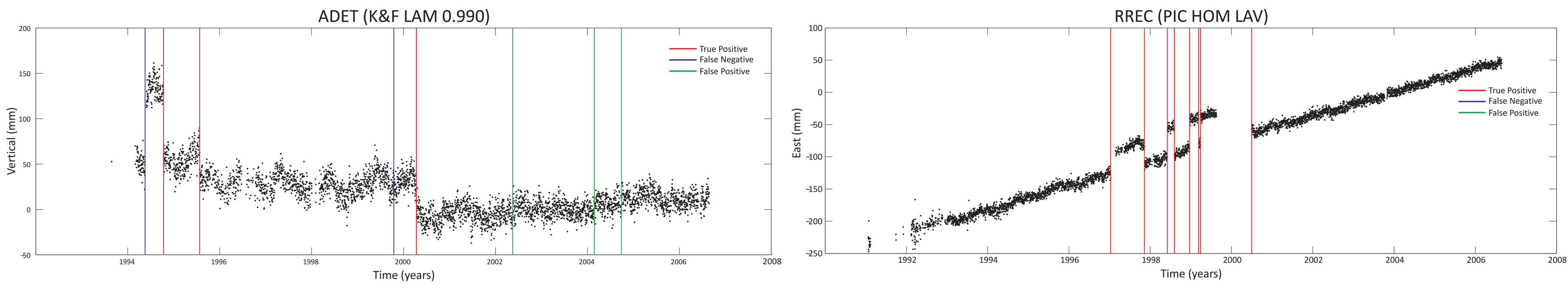
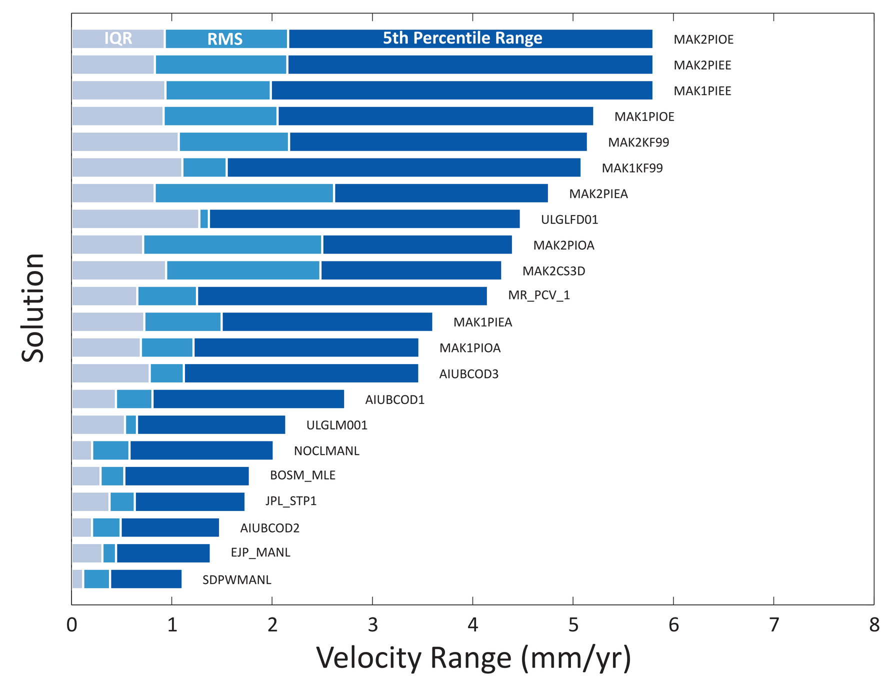


Figure 2 Examples of two sites showing success/failure using the (right) Picard (**PIC**) approach with **HOM**oscedastic assumption and the **LAV**ielle (2005) penalty function (**PIOA**) and (left) the Kehagia and Fortin (**K&F**) with a scaling parameter set to 0.99 (**KF99**)

Acknowledgements

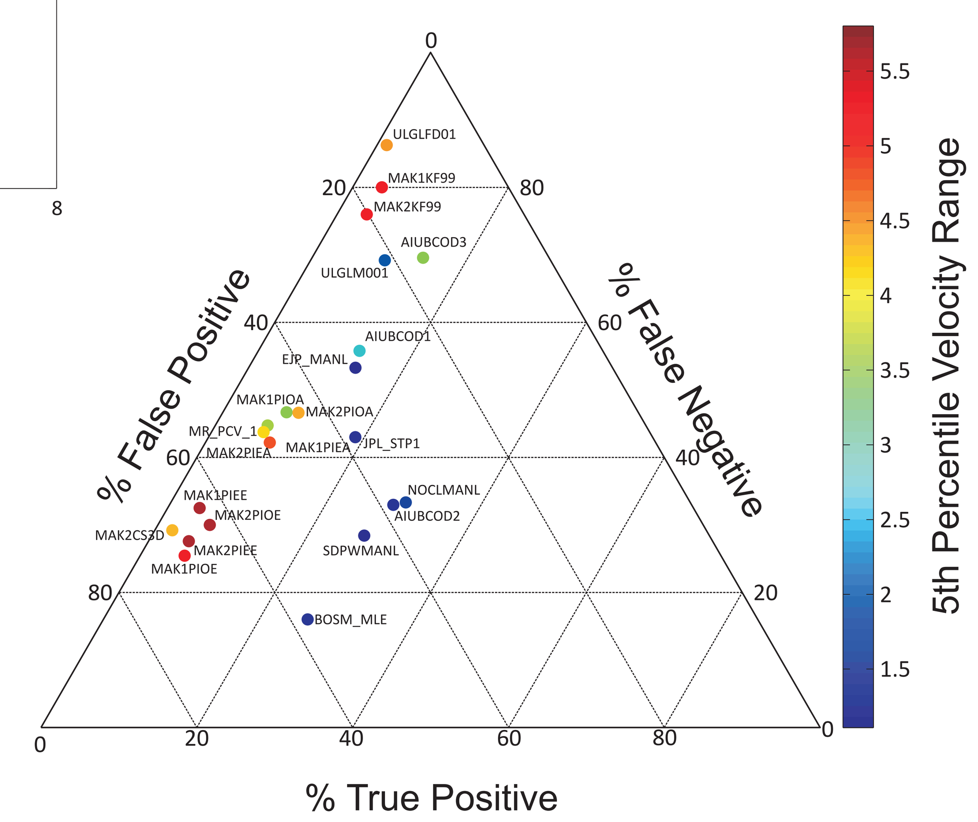
This research was supported by NERC grants and COST Action ES0701. We thank Olivier Mestre for helpful discussions and the solutions providers (shown above)

Results



We chose to present the 5th percentile ranges (5% to 95%) in velocity differences (from truth) as a metric for the performance of the different solutions. Figure 3 (left) ranks the solutions in order of their performance in this metric. Also shown are the RMS and the Interquartile Range (25% to 75%). The lower the 5th percentile range, the closer the solution is to the truth.

Figure 4 (right) depicts the ratios of the three variables, TP, FP and FN as positions in an equilateral triangle together with their 5th percentile velocity range. A zero velocity range would occur in the bottom right hand corner of the triangle. The plot highlights the trade off between the three. Over segmentation leads to a higher FN %, for example the KF99 solution. The manual solutions have a low FN % but a higher FP rate compared to JPL_STP1 and AIUBCOD2.



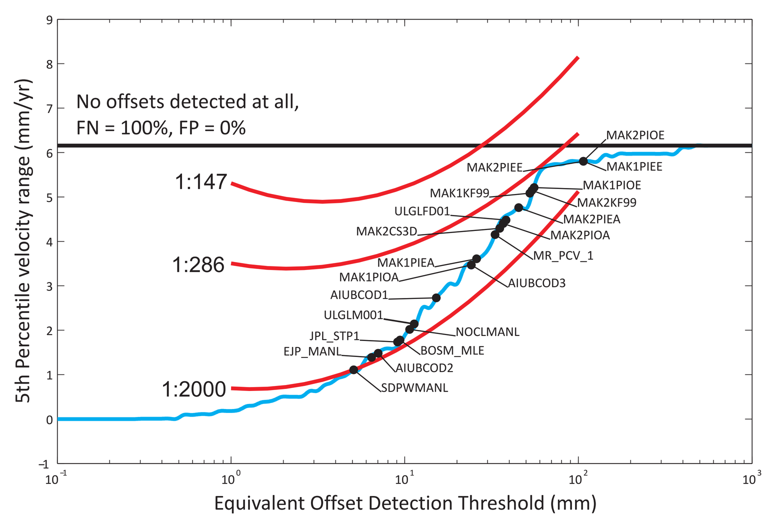
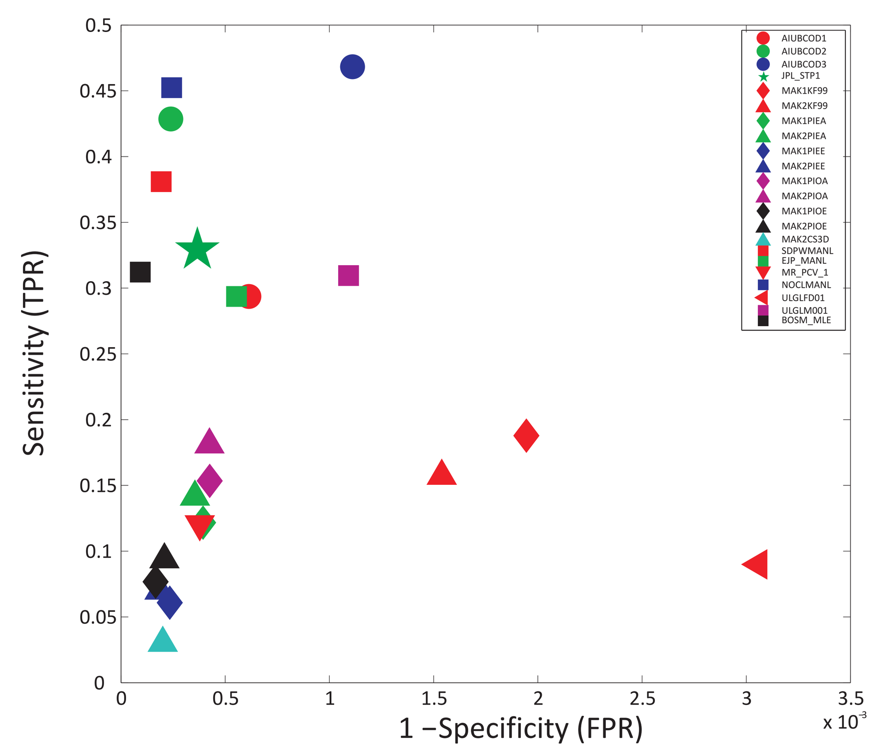


Figure 5 (left) shows the results of some synthetic runs. Blue line has no FP but only offsets greater than the detection threshold are set as TP's. The 5th percentile range estimated gives a rough idea of what size offsets are being detected by the solutions. For instance JPL_STP1 has an equivalent offset detection threshold of around 10 mm. Red lines are a similar simulation with added FP's with a ratio of total offsets (TP + FP + FN) to timeseries length given next to the lines. Too much segmentation can lead to velocity biases that are larger on average than ignoring all offsets!

Receiver Operator Characteristic (ROC) curve for the 11 methods (22 solutions) used so far. An ideal method would have a high sensitivity (true positive rate) and high specificity (low false positive rate) : i.e. top left hand corner

TPR = TP / (TP + FN)

FPR = FP / (FP + TN)



Conclusions

Further work is required in order to accurately determine offsets in GPS time series and produce more sophisticated metrics for classifying the solutions. Over-segmentation and mis-identification of offsets near the ends of time series remain problematic. It is likely that techniques developed in other fields will not provide an optimal solution for GPS, at least without modification. We will be looking into Bayesian statistics for more robust classification of the ratios of TP, FP and FN in each solution. We encourage individuals and groups to download the DOGEx dataset and provide solutions to us for assessment.

References

Kehagias, A., and V. Fortin. 2006. Time series segmentation with shifting means hidden markov models, Nonlin. Processes Geophys., 13 (3), 339-352

Lavielle, M. 2005. Using penalized contrasts for the change-point problem, Signal Processing, 85 (8), 1501-1510, 10.1016/j.sigpro.2005.01.012.

Lebarbier, E. 2005. Detecting multiple change-points in the mean of Gaussian process by model selection, Signal Processing, 85 (4), 717-736, 10.1016/j.sigpro.2004.11.012.

Picard, F., S. Robin, M. Lavielle, C. Vaisse, and J.J. Daudin. 2005. A statistical approach for array CGH data analysis, Bmc Bioinformatics, 6, 10.1186/1471-2105-6-27.

NATURAL ENVIRONMENT RESEARCH COUNCIL

COST ACTION ES0701 improved constraints on models of GIA