EU-VGOS activities in Vienna

F. Jaron, S. Bernhart, J. Böhm, J. González García, J. Gruber, Y. K. Choi, I. Martí-Vidal, M. Schartner, B. Soja, J. Wagner, E. Varenius, H. Verkouter, on behalf of the EU-VGOS collaboration

Abstract The EU-VGOS collaboration carries out observational sessions on a regular basis, some of them being correlated at TU Wien. This is a progress report on aspects of data processing from e-transfer to fringe-fitting. We compare two approaches of handling dual linear polarization: combining all four linear polarization products to pseudo Stokes *I*, and conversion from linear to circular polarization using PolConvert. Our results so far seem to indicate that the latter approach gives a higher signal to noise ratio than the former.

Frédéric Jaron · Johannes Böhm · Jakob Gruber
Department of Geodesy and Geoinformation, TU Wien, Wiedner Hauptstraße 8, A-1040 Vienna, Austria
`frederic.jaron@tuwien.ac.at`

Simone Bernhart · Yoon Kyung Choi
Reichert GmbH, Hittorfstraße 26, DE-53129 Bonn, Germany

Javier González García
Observatorio de Yebes (IGN), Apartado 148, 19181, Yebes, Spain

Ivan Martí-Vidal
Departament d'Astronomia i Astrofísica, Universitat de València, C. Dr. Moliner 50, 46100, Burjassot, València, Spain; Observatori Astronòmic, Universitat de València, Parc Científic, C. Catedràtico José Beltrán 2, 46980, Paterna, València, Spain

Matthias Schartner · Benedikt Soja
Institute of Geodesy and Photogrammetry, ETH Zürich, 8093, Zurich, Switzerland

Jan Wagner · Frédéric Jaron
Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany

Eskil Varenius
Department of Space, Earth and Environment, Chalmers University of Technology, Onsala Space Observatory, SE-439 92, Onsala, Sweden

Harro Verkouter
Joint Institute for VLBI ERIC, Dwingeloo, Netherlands

Keywords VGOS, E-transfer, Correlation, Fringe-fitting, Polarization

1 Introduction

The EU-VGOS project was started in 2018 by Alef et al. (2019) with the aim to find a way for post-processing VGOS data alternate to the way implemented by the Haystack group within the *HOPS* package (Niell et al., 2018, and references therein). The idea is to exploit a *CASA* (McMullin et al., 2007) based *Python* pipeline that includes the *PolConvert* (Martí-Vidal et al., 2016) software, makes use of phase-cal tones, is able to correct ionospheric dispersion, and allows Stokes *I* fringe-fitting of the data. For this purpose, the EU-VGOS collaboration carries out research and development sessions with a network of VGOS stations, mainly located in Europe but also beyond. Here we report on the processing of EU-VGOS data at the VLBI correlator of TU Wien, from e-transfer to fringe-fitting.

The TU Wien VLBI correlator is operated by the working group of Higher Geodesy at the Department of Geodesy and Geoinformation. Correlations are being carried out by J. Gruber and F. Jaron. For data storage and processing we use the infrastructure of the Vienna Scientific Cluster[1], where we can use up to ten compute nodes simultaneously. Each of these nodes is equipped with 2 CPUs with 24 cores each, resulting in 48 cores per node. We have 1 PB of disk space reserved exclusively for VLBI. We are connected to the GÉANT network via a 10 Gbit/s link which is, however, shared with other users. For correlation we use the

[1] `http://www.vsc.ac.at`

DiFX-2 software correlator (Deller et al., 2011) in various sub-versions. For fringe-fitting we currently use *HOPS/fourfit* in its version 3.22. We have also *PIMA* (Petrov et al., 2011) installed on our system.

2 Short-scan experiments

Here we report on the processing of five[2] EU-VGOS sessions: ev0287, ev0303, ev0314, ev0328, ev0345. These experiments were scheduled by M. Schartner using his software *VieSched*++ (Schartner & Böhm, 2019). One special feature of the sessions is that shorter integration times were tested. Currently, all IVS-VGOS sessions utilize a fixed 30 s integration time, independent of the source brightness. In contrast, within the sessions considered here, integration times between 10 and 20 s were examined. Therefore, each session was divided into blocks and each block was scheduled with a different integration time. In addition, calibrator scans were included with durations of 2 to 3 min each. To ensure a reasonable signal to noise ratio (SNR) for the short scans, the theoretical SNR was computed during scheduling. Although this is standard for S/X observations, it is currently not utilized in the IVS-VGOS sessions due to the lack of station system equivalent flux density (SEFD) and source flux density information in the scheduling catalogs for the VGOS frequencies. To overcome this issue, the SEFD information was directly provided by the stations and the source flux densities were interpolated and extrapolated based on their S and X band flux densities. Similar to scheduling standard S/X observations, the SNRs were computed for each band individually and different thresholds between 8 and 20 per band were used. In these five sessions the participating stations were the two Onsala telescopes (ONSA13NE and -SW), Wettzell (WETTZ13S), Yebes (RAEGYEB), and in ev0345 also Ishioka (ISHIOKA) in Japan.

The frequency setup of these sessions (Fig. 1) is identical to the one currently used for global VGOS experiments, with four bands distributed between 3 and 11 GHz. The observations, each six hours in duration, were carried out in dual linear polarization. The polarizations were labeled "X" and "Y" in the VEX files. It is important to point out that these are local horizontals/verticals at the telescopes and (without a field rotator) are not aligned with the celestial *X* and *Y* coordinates, and are subject to changes in the parallactic angle. Hence, "H" and "V" might be a better choice for these labels. The sampling rate per channel was 64 Msamples/s (and 32 Msamples/s for Yebes I/Q sampling) with 2-bit quantization, resulting in a net data rate of 8.192 Gbit/s per antenna. All stations are equipped with phase-cal units, most of them inject tones every 5 MHz, except for Yebes, which has tones every 10 MHz. Most of the stations have DBBC backends (Tuccari et al., 2018), providing real-valued data, only Yebes has an RDBE (Ruszczyk et al., 2012) and the data are complex-valued. Stations record their data onto Mark 6 modules or onto FlexBuff recorders, except for ISHIOKA, who use an ADS3000+ sampler and a K5VSI recording system.

[2] This number is continuously growing. Currently ten of these sessions are being processed in Vienna, but we restrict ourselves to these five experiments as representative examples.

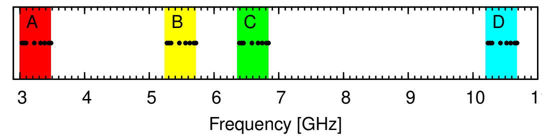

Fig. 1 Frequency setup. Each of the four bands (A-D) has a width of 480 MHz and contains 8 channels, each 32 MHz wide.

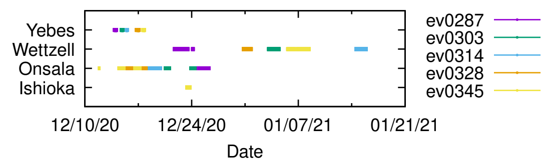

Fig. 2 E-transfer of five EU-VGOS sessions.

3 E-transfer

The TU Wien correlator does not have any Mark 6 units at its disposal, so data have to be transferred electronically. We used *m5copy* and *jive5ab* for the data transfers[3]. It took approximately one month to complete all transfers with a total volume of 115 TB (Fig. 2). When comparing the transfer duration of the individual stations, one has to keep in mind that Onsala had to transfer twice the amount of data (having two VGOS tele-

[3] `https://github.com/jive-vlbi/jive5ab`

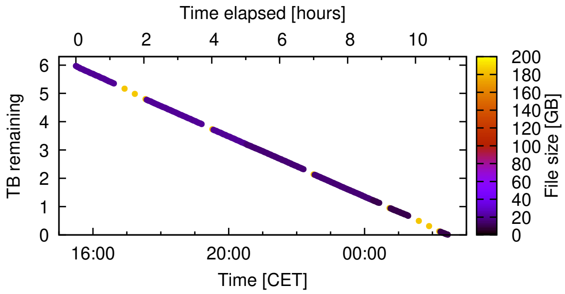

Fig. 3 E-transfer of ev0303 from Yebes to Vienna. Remaining data volume is plotted against time. The file size is color coded.

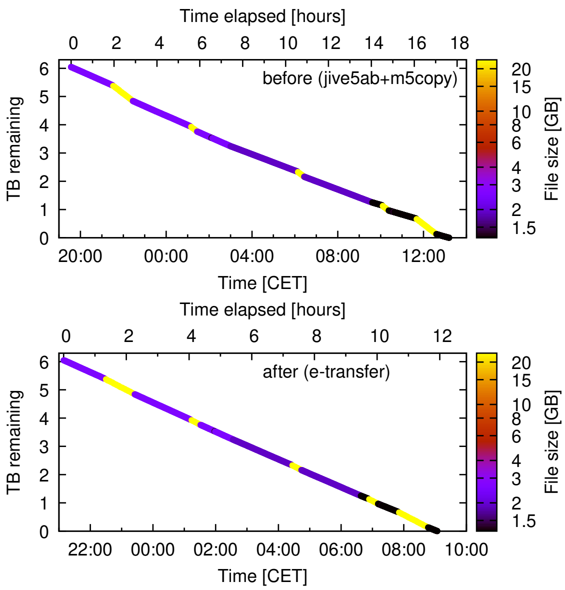

Fig. 4 E-transfer of ev0303 from Onsala to Vienna. **Top:** There is a strong dependence of transfer rate on file size. **Bottom:** Using the e-transfer daemon/client pair almost completely removes the file size dependence.

scopes), and Ishioka participated in only one (ev0345) out of these five sessions.

As an example of a smooth transfer, Fig. 3 shows a plot for Yebes. The curve of remaining data volume plotted against time is a straight line without any dependence on file size. The average transfer rate in this case was 1.2 Gbit/s. The situation was initially very different for the transfers from Onsala. The plot in the top panel of Fig. 4 reveals a strong dependence of transfer rate on file size. Onsala recorded the data in multi-file format, resulting in many files with relatively small sizes of a few GB each. Only the few larger calibrator scans reached the full speed of 1.3 Gbit/s. The small

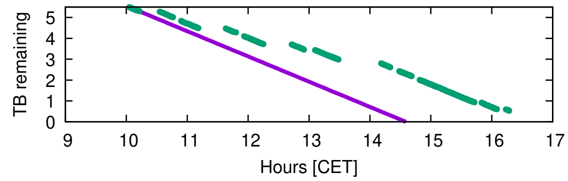

Fig. 5 Conversion from multi- to single-thread using *vmux* on Wettzell data of session ev0303. Remaining data volume is plotted against time. Gaps that follow the linear trend correspond to larger files being converted. Horizontal gaps are related to *vmux* being stuck in an infinite loop of error messages. The purple line is an extrapolation of the initial linear trend that would have been continued in the absence of errors.

files came at only half that speed. After discussing the issue with the author of *jive5ab* it became clear this is the result of a large-ish constant per-file delay between transfers that the *jive5ab+m5copy* approach suffers from, causing the experienced transmission time of a small file to be (significantly) larger than the physical transfer time of the content itself. H. Verkouter suggested trying the e-transfer daemon/client pair[4] – a server/client file transfer utility that, like *jive5ab* but without its overheads, is capable of using the UDT data transfer protocol (Gu & Grossman, 2007) for fast data transfer over long, fat links where the TCP protocol significantly underperforms. We tested the new toolchain by repeating the same transfer. The resulting plot (bottom panel of Fig. 4) confirms that the file size dependence has almost been removed. In this way the total transfer duration could be reduced from initially 18 down to 12 hours.

4 Correlation

4.1 Converting multi- to single-thread

The *DiFX* software correlator adopted also for VGOS can process recordings that are multi-threaded VDIF (e.g., from the VLBA) or multi-threaded multi-datastream VDIF (e.g., Phased NOEMA in the EHT and GMVA). However, only single-threaded VDIF and the older *DiFX* version 2.5.3 have been validated for VGOS so far. In addition, the *difx2mark4* converter in 2.5.3 and in versions prior to 2.6.2 does not correctly handle PCal entries from multi-datastream recordings.

[4] https://github.com/jive-vlbi/etransfer

Thus any recordings not present in a VGOS-validated layout should be converted prior to correlation in *DiFX*-2.5.3.

For the five sessions that are subject of this article, Wettzell recorded the data in multi-threaded VDIF format. Conversion is done using the tool *vmux* that is part of the *DiFX* distribution[5]. Unfortunately, converting the data took very long, as shown in Fig. 5. The conversion of six hours of data from one station alone took approximately as long as the observation. This, however, includes time during which *vmux* was outputting error messages[6] in an infinite loop which had to be manually interrupted. Another consequence of these errors is that about 5 % of the data could not be converted, with calibrator scans being especially affected.

We have by now implemented a script which automatically interrupts and restarts the process when such error messages occur, eliminating this overhead. Nevertheless, *vmux* still takes very long compared to the actual observing time and will dominate the processing time if many stations deliver the data in multi-thread format. Attempts to parallelize the process on a cluster have not been successful so far. One possible reason could be that the process is I/O-limited.

4.2 DiFX for VGOS

Currently there has not been any official *DiFX* release which correlates VGOS data without any problems. Correlator centers have to apply various patches to their installations. Aiming at a *DiFX* version for VGOS, we discussed the issue within the EU-VGOS collaboration. As a result, at MPIfR Bonn we backported a specific set of EU-VGOS-relevant patches from the more recent *DiFX*-2.6 series into the *DiFX*-2.5 development base. In addition, MIT Haystack updated the included *HOPS* package from 3.19 to the current 3.22 vendor release used in VGOS processing. This development base, including HOPS 3.22, will be released as *DiFX*-2.5.4 after completion of an open testing phase in the *DiFX* user community[7]. At Vienna we used this version to correlate ev0303, and the results are shown in the following.

[5] If you have Mark 6 units, another possibility is to use the tool *mk6vmux* to convert the data directly off Mark 6 modules.

[6] Messages of the type: “desync, frame delta 15796, fps 31250”

[7] `https://deki.mpifr-bonn.mpg.de/Cooperations/EU-VGOS_P-o-C/VGOS_DiFX_Versions`

5 Calibration and fringe-fitting

We focus on two options for fringe-fitting:

1. Combine the four linear polarization products to pseudo Stokes *I* without any actual polarization calibration (cf. Eq. 2 in Cappallo 2014). This is implemented in the *HOPS/fourfit* software suite and is described in the MIT Haystack VGOS manual[8].

2. Convert visibilities from linear to a circular polarization basis, including a proper cross-polarization complex gain calibration (Martí-Vidal et al., 2016). Using a preliminary version of *PolConvert* for VGOS, we estimate the cross-polarization gain in a least squares method and apply the estimates in the conversion process. *PolConvert* also takes care of the differential parallactic angle correction to be introduced between polarization products in the visibilities. After this, we are ready to tell fourfit[9] to combine the two parallel hand circular polarization products to form true Stokes *I*.

6 Results

Figure 6 shows two *fourfit* fringe-plots for one calibrator scan from ev0303. The plot in the top panel shows the result for pseudo Stokes *I*, and the bottom panel for the polconverted data. For this example scan pseudo Stokes *I* yields an SNR of 525.4 while the polconverted data give a considerably higher SNR of 607.7. An inspection of other scans confirmed a tendency of higher SNR for polconverted data.

7 Conclusions and outlook

We have presented a work in progress. These are our conclusions so far:

1. E-transfer to Vienna works smoothly. An initial drop in transfer rate for small files could be fixed by using a different file transfer utility.

[8] `https://www.haystack.mit.edu/wp-content/uploads/2020/07/docs_hops_000_vgos-data-processing.pdf`

[9] Command line option -P RR+LL

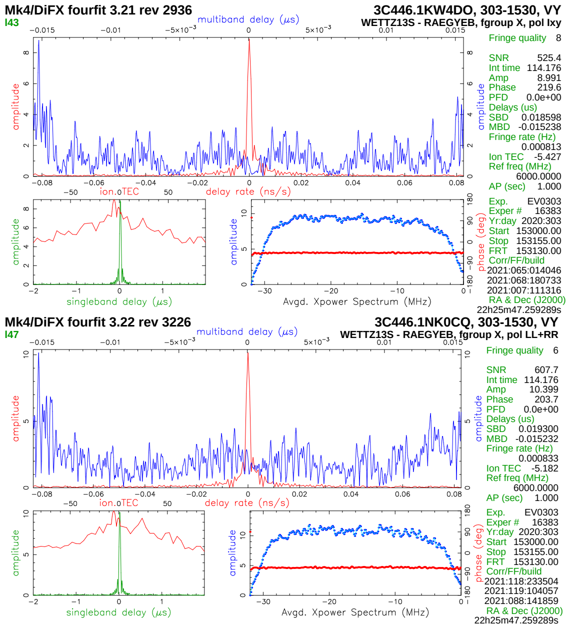

Fig. 6 Fourfit fringe-plot. **Top:** Pseudo Stokes *I*. **Bottom:** Polconverted data.

2. Converting multi-thread VDIF data to single-thread, using the tool *vmux*, takes very long and will become a significant overhead for experiments with larger networks.
3. In order to avoid locally patched *DiFX* installations for VGOS, Bonn and MIT Haystack have put together *DiFX*-2.5.0. This version is in a testing phase now and needs to be approved by the *DiFX* and VGOS communities to become version 2.5.4.
4. First results from fringe-fitting polconverted data indicate that SNR seems to be systematically higher than for pseudo Stokes *I*. This is expected, because *PolConvert* does a proper cross-polarization complex gain calibration.

Until now, the gain estimates in *PolConvert* come from only one calibrator scan selected from the whole experiment. Further development of the tool has already shown that many scans can be used in the least squares adjustment to generate more accurate solutions, and the current status of the development is concentrated on being able to apply this functionality to any of the EU-VGOS session setups.

Algorithms for fringe-fitting dual linear polarization broadband data have now been implemented in PIMA. We plan to include this option in the future.

Acknowledgements

We thank Alan Roy of the MPIfR Bonn for carefully reading the manuscript and providing useful comments. The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC). This research was funded in part by the Austrian Science Fund (FWF) [P31625].

References

Alef W, Anderson JM, Bernhart S et al. (2015) The EUropean-VGOS Project. *Proceedings of the 24th European VLBI Group for Geodesy and Astrometry Working Meeting*, pp. 107-111

Deller AT, Brisken WF, Phillips CJ et al. (2011) DiFX-2: A More Flexible, Efficient, Robust, and Powerful Software Correlator. *PASP*, 123, 901, pp. 2752, doi:10.1086/658907

Cappallo R (2014) Correlating and Fringe-fitting Broadband VGOS Data. *IVS 2014 General Meeting Proceedings*, p. 91-96, ISBN 978-7-03-042974-2

Gu Y, Grossman RL (2007) UDT: UDP-based data transfer for high-speed wide area networks. *Computer Networks*, 51, 7, p. 1777-1799, doi:10.1016/j.comnet.2006.11.009

Martí-Vidal I, Roy A, Conway J, Zensus JA (2016) Calibration of mixed-polarization interferometric observations. *A&A*, 587, A143, 8 pp., doi:10.1051/0004-6361/201526063

McMullin JP, Waters B, Schiebel D et al. (2007) CASA Architecture and Applications. *Astronomical Data Analysis Software and Systems XVI ASP Conference Series*, 376, p.127

Niell A, Barrett J, Burns A et al. (2018) Demonstration of a Broadband Very Long Baseline Interferometer System: A New Instrument for High-Precision Space Geodesy. *Radio Science*, 53, 10, pp. 1269-1291, doi:10.1029/2018RS006617

Petrov L, Kovalev YY, Fomalont EB, Gordon D (2011) The Very Long Baseline Array Galactic Plane Survey – VGaPS. *AJ*, 142, 2, 35, 23 pp., doi:10.1088/0004-6256/142/2/35

Ruszczyk C, McWhirter R, Beaudoin C et al. (2012) VLBI2010 using the RDBE and Mark 5C. *Seventh General Meeting (GM2012) of the international VLBI Service for Geodesy and Astrometry (IVS)*, p. 81-85

Schartner M, Böhm J (2019) VieSched++: A New VLBI Scheduling Software for Geodesy and Astrometry. *PASP*, 131, 1002, pp. 084501, doi:10.1088/1538-3873/ab1820

Tuccari G, Alef W, Dornbusch S et al. (2018) DBBC3 the new wide-band backend for VLBI. *14th European VLBI Network Symposium & Users Meeting*. Online at `https://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=344`, id.140