Tycho-Gaia and beyond

or: What Daniel researched in his Ph.D. thesis
or: Why the first Gaia data release is special

D. Michalik (Lund University)

2010–2015: Ph.D. at Lund University, Sweden
2009–2010: Young Graduate Trainee, European Space Agency

Step 1

Introduction

The first Gaia data release

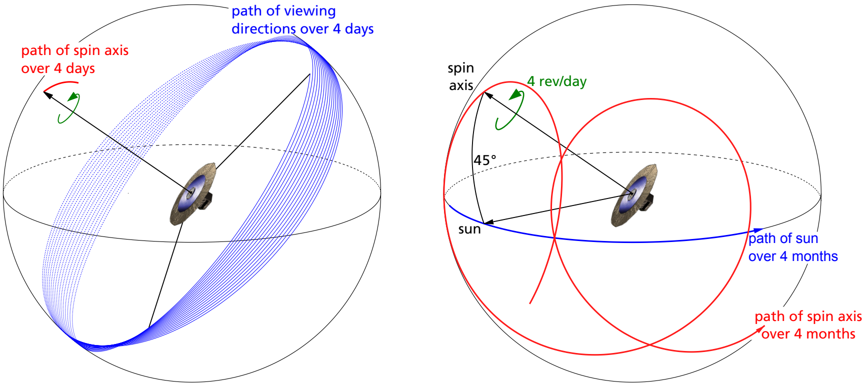
Release schedule revised based on Ph.D. thesis research

Current prediction

- Summer 2016:
  - **Improved positions, realistic uncertainties and G-magnitudes for 100s of millions of stars**
  - Ecliptic pole data (photometry calibration)
  - **5 parameter astrometry for approx. 2 million Tycho and Hipparcos stars (Tycho-Gaia Astrometric Solution; TGAS)**

Full five parameter Gaia-only astrometry from summer 2017

Gaia scanning the sky



Figures: L. Lindegren & D. Michalik

Gaia scanning the sky

Sky coverage over time ($Nsl37 + Nsl_{GAREQ}$)

Animation: D. Michalik & B. Holl

Gaia data processing

Preprocessed observations ⇒ Individual source parameters

- Have: 10^{12} observations
- Want: ca $5x10^9$ unknowns:
  Astrometric source parameters, attitude, calibration
- How: **Globally**, self-consistent manner
- Strategy: **Iterative Solution**
- Tool: Astrometric Global Iterative Solution (AGIS)

⇒ A linear least-squares problem $\boldsymbol{N}\boldsymbol{x} = \boldsymbol{b}$ with iterative solution

Read more: Lindegren et al. (2012, A&A)

AGISLab

- For development and testing of the AGIS algorithms
- For experiments with scientific exploitation

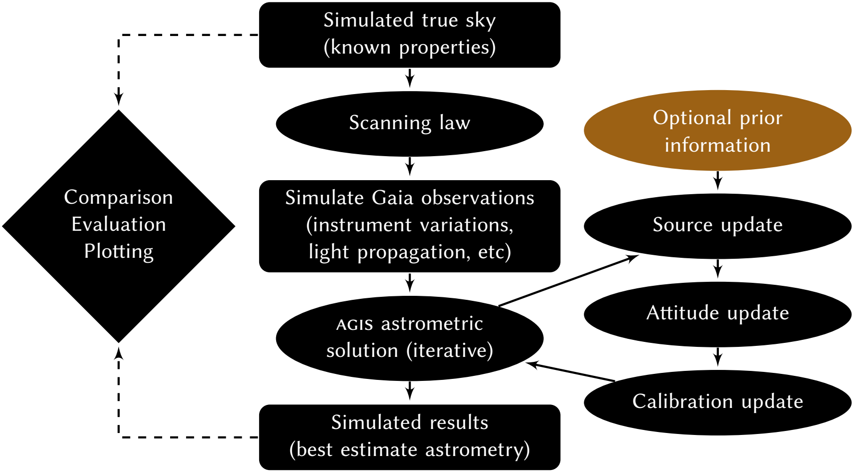
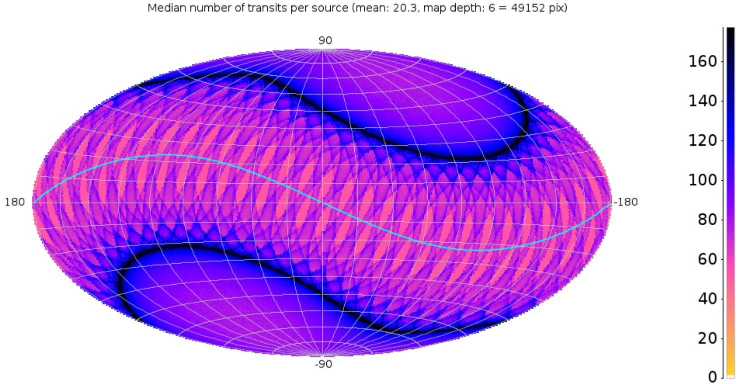


Figure: D. Michalik

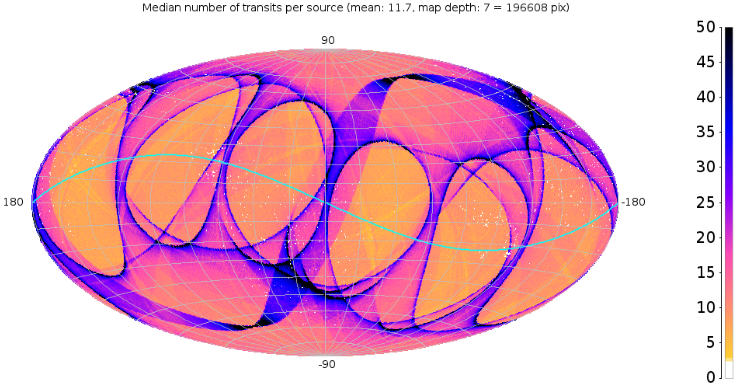
First release is special



Number of transits in a nominal 5 year interval: smooth coverage, 80 transits on average

Simulations: D. Michalik

First release is special



Number of transits during the 13 months for Gaia-DR1: some areas are poorly observed

Simulations: D. Michalik

Step 2

Ambiguity in early datasets

Change in observed coordinate over 5 years

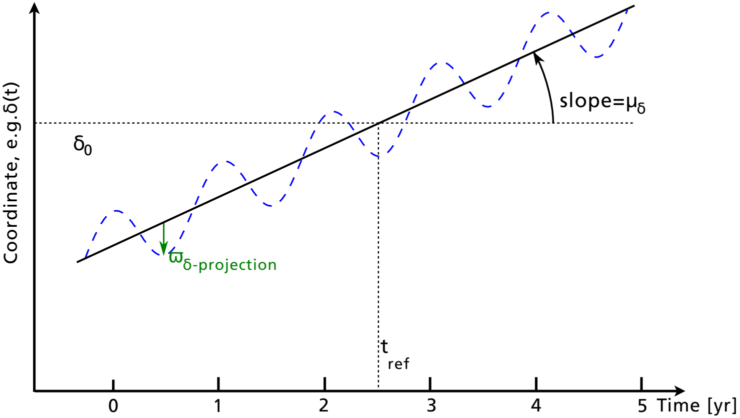


Figure: D. Michalik & L. Lindegren

Gaia observations over 5 years ⇒ pos, ϖ, μ

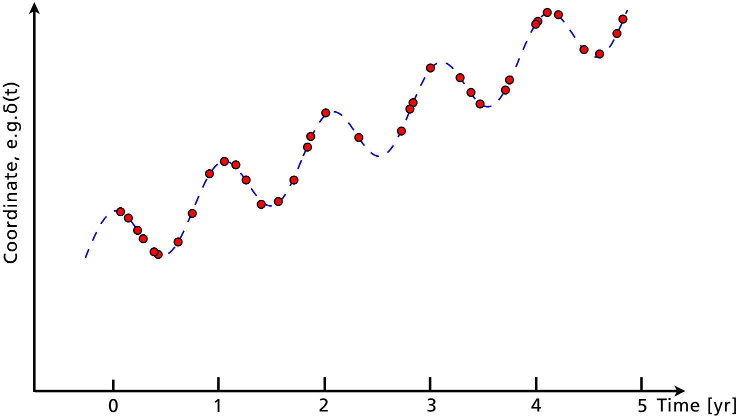


Figure: L. Lindegren & D. Michalik

Gaia observations over 1 year ⇒ marginal

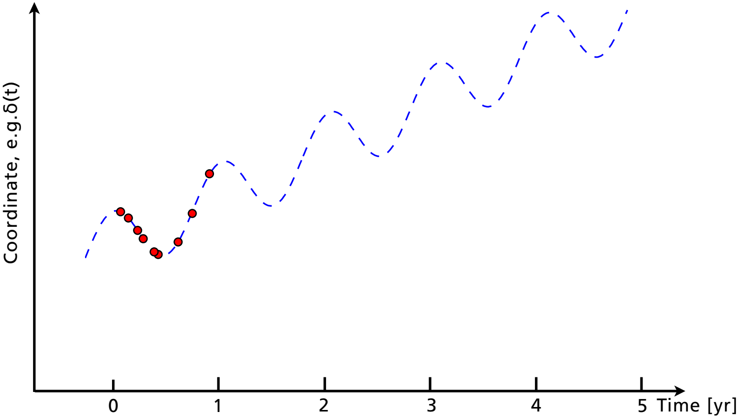


Figure: L. Lindegren & D. Michalik

$\mu - \varpi$ degeneracy for < 1 year observations

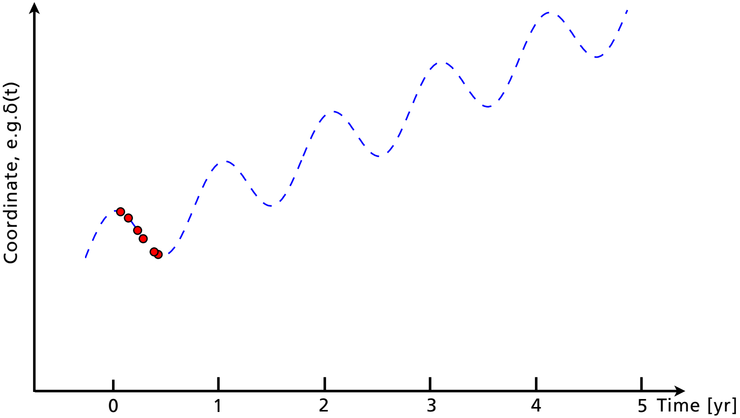


Figure: L. Lindegren & D. Michalik

$\mu - \varpi$ degeneracy for < 1 year observations

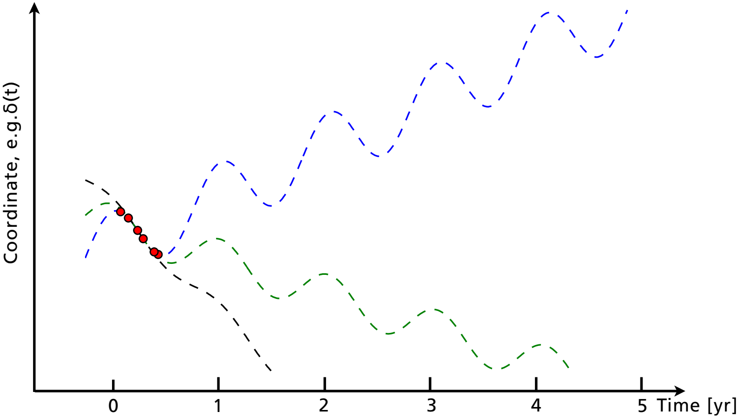


Figure: L. Lindegren & D. Michalik

Step 3

Prior information to the rescue

Integrating prior data in Gaia astrometry

Using Bayes' rule $f(\boldsymbol{x}|\boldsymbol{h}) \propto L(\boldsymbol{x}|\boldsymbol{h}) \times p(\boldsymbol{x})$

- Prior probability density function $p(\boldsymbol{x})$ from prior data
- Likelihood $L(\boldsymbol{x}|\boldsymbol{h})$ from Gaia
- Assuming Gaussian errors: posterior $f(\boldsymbol{x}|\boldsymbol{h})$ is given by joint solution of combined normal equations

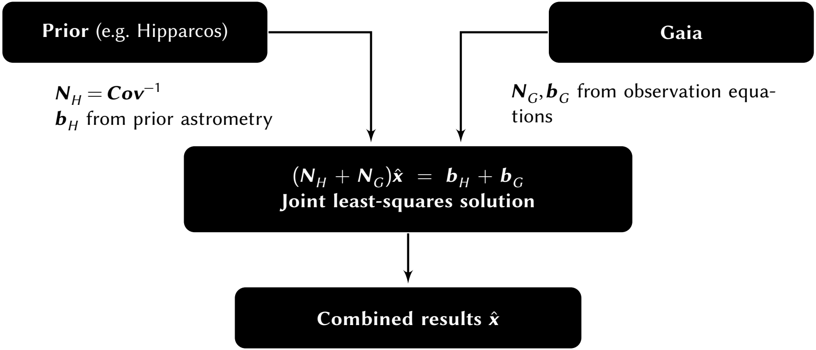


Figure: D. Michalik

Step 4

Tycho-Gaia Astrometric Solution

The Tycho-2 catalogue

2.5 million positions at J1991.25, σ = 5–70 mas, 90% complete to V=11.5, obtained from Hipparcos starmapper[1]

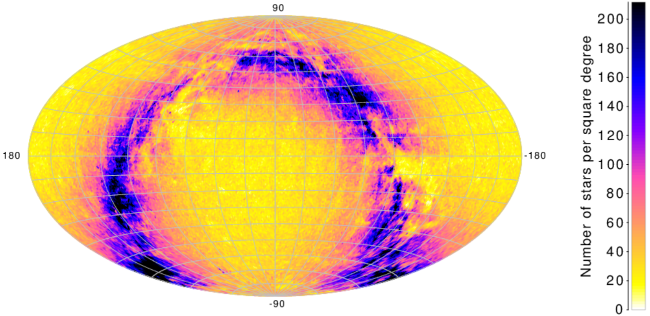


Figure: Tycho-2 sky coverage (Michalik et al. 2015a, Fig. 2, left)

[1] auxiliary photomultiplier and grid for attitude determination

Position alone sufficient to lift the degeneracy

⇒ **Independent** long-baseline proper motions, parallaxes

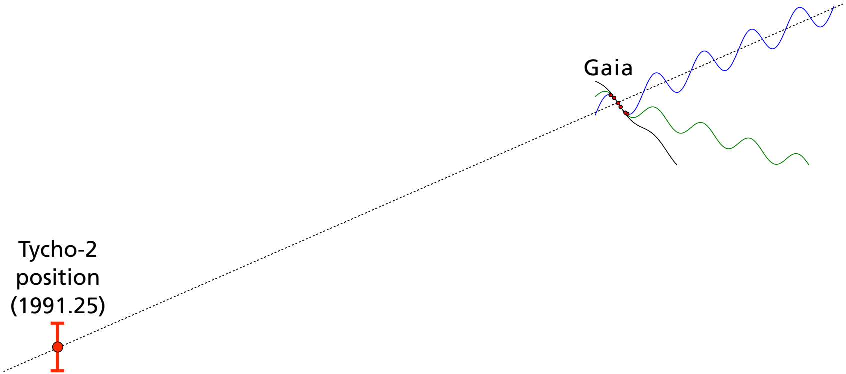
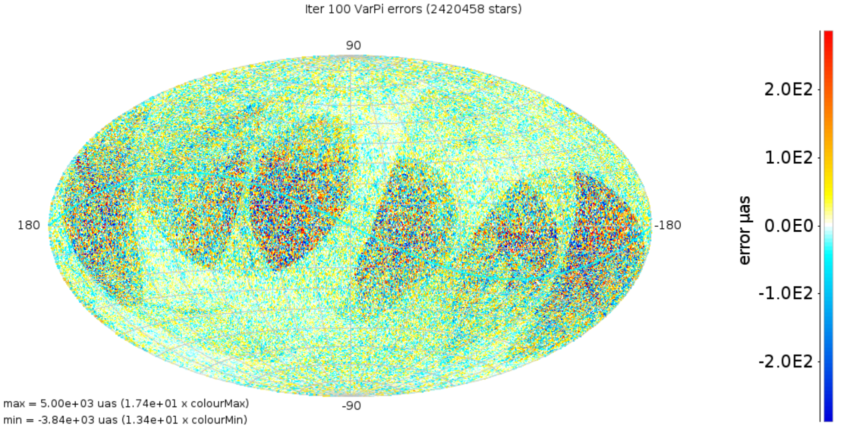


Figure: L. Lindegren & D. Michalik

Simulated Gaia observations (July 2014–May 2015)

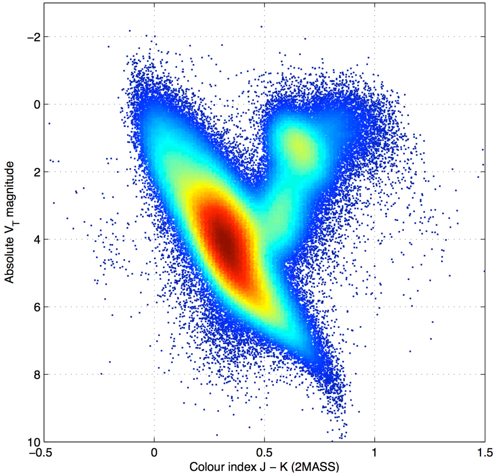


Unbiased parallax errors (average per pixel), small spread
⇒ **Success!** (...in a perfect world)

Figure: D. Michalik

HR-diagram from TGAS trial (real data)

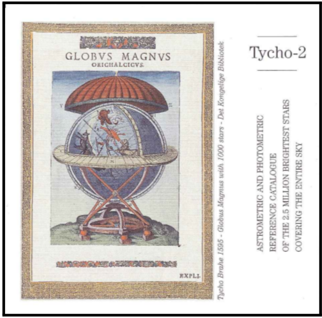
Results are based on a trial run using just a few months of Gaia data

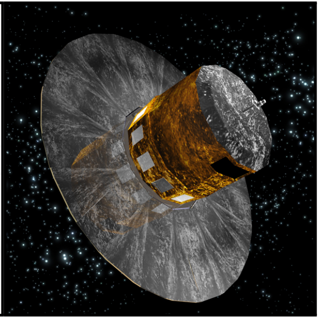


(approx. 480 000 Tycho-2 stars, with 2MASS colours, $\varpi > 0$, $\sigma < 1$ mas, $\varpi/\sigma > 10$)

Plot: L. Lindegren

Tycho-Gaia Astrometric Solution (TGAS)





- Prior: add positions at J1991.25 as additional observations
- Full solution with much less Gaia data (approx. one year earlier)
- Hipparcos stars are an integral part of TGAS
- **Independent proper motions and parallaxes for 2.5 M stars**

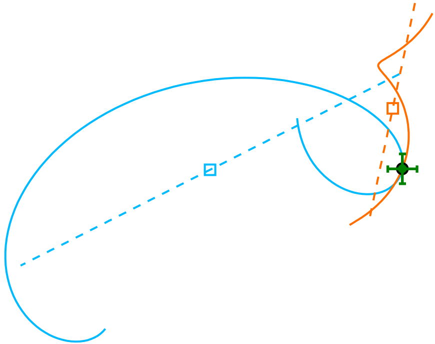
Left: Tycho-2 CD cover
Right: Illustration Gaia satellite (ESA)
Read more: Michalik et al. 2014, 2015a

Step 5

A prior to rule them all

Short/scarce datasets need a prior

- Gaia observes 1 billion of the few 100 billion Milky Way stars
- Tycho prior for 2.5 *M* – what about the remaining 997.5 *M*?
- Actual parallax and proper motion cause (unknown!) bias



- Same partial solution for very different astrometric parameters
- Is the observation a nearby dwarf (blue) or a distant giant (orange)?
- Formal errors grossly underestimate the actual errors!

Figure: L. Lindegren & D. Michalik

A generic approach for incomplete data

Objective: Obtain positions and correct error estimates, even for:

- First release (too short a time interval)
- Stars at the detection limit (seen too seldom)
- Transient objects (too short a time interval)

What is ...

- the influence of a prior to an astrometric solution?
- the probability density function of the positional offset?
- a realistic distribution of true ϖ and μ?
- the optimal prior to pick, and what does it depend on?

Study based on Gaia Universe Model Snapshot (GUMS)

Read more: Michalik et al. 2015b

Behaviour of astrometric solution with prior

Left: Quasi two parameter solution

- Formal errors grossly underestimate actual errors

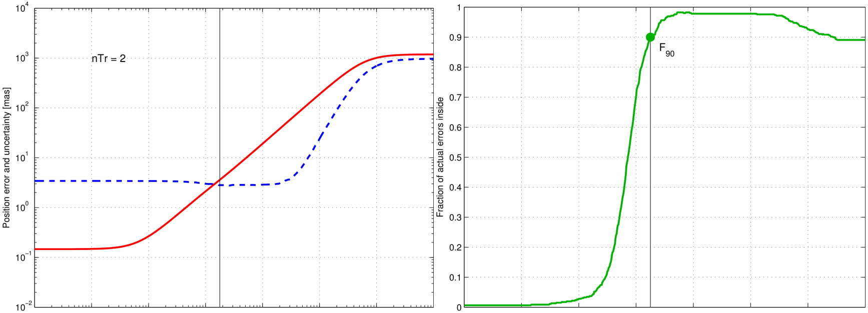
Middle: Use knowledge that parallaxes, proper motions are small

- 5 parameter solution, realistic formal errors

Right: Degenerate solution

Figures: Michalik et al. 2015b, Fig. 1+2

Generic prior properties

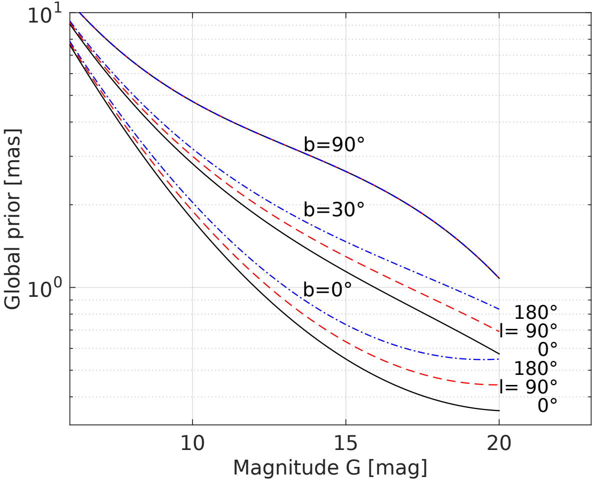


90% of the actual position errors contained in the 90% confidence formal uncertainty ellipse

Excerpt from Michalik et al. 2015b, Fig. 3

Generic prior properties

Prior uncertainty depends on magnitude and Galactic latitude



Generic prior results

Table: Actual errors and agreement factor with formal uncertainty.

Prior $\sigma_{\varpi,p}$	Fraction in 90% conf. ellipse			Actual position errors [mas]		
	$G \simeq 11$	$G \simeq 15$	$G \simeq 19$	$G \simeq 11$	$G \simeq 15$	$G \simeq 19$
none (2 parameters)	0.5%	1.8%	13.5%	33.0	16.3	15.2
Generic prior	90.1%	91.4%	91.2%	7.6	4.3	7.6

Benefits: always provides a non-singular solution

1. Reasonable error estimates and better actual errors
2. With insufficient amount of observations

Caveats: biases the solution

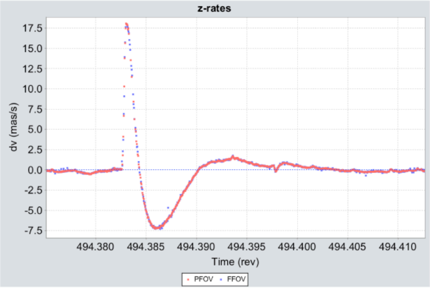
1. Serious for $\varpi, \mu \Rightarrow$ Not to be published
2. Must not be used as soon as enough information are available

Table: Excerpt from Michalik et al. 2015b, Table 2

Step 6

Verification of TGAS parallaxes

Real life ≠ simulations



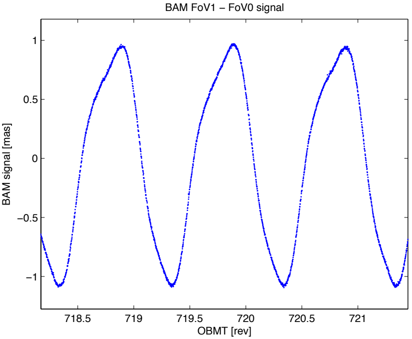
Some of the real life complications:

- Data gaps due to orbit maintenance, cosmic rays
- Transmission loss ⇒ Heating for decontamination
- Re-focussing
- Micro-meteoroid hits
- Thermal micro-clanks (material relaxation)

And the eternal nightmare of an astrometrist ...

Figure: a real Gaia micro-meteoroid hit

Basic Angle Variations

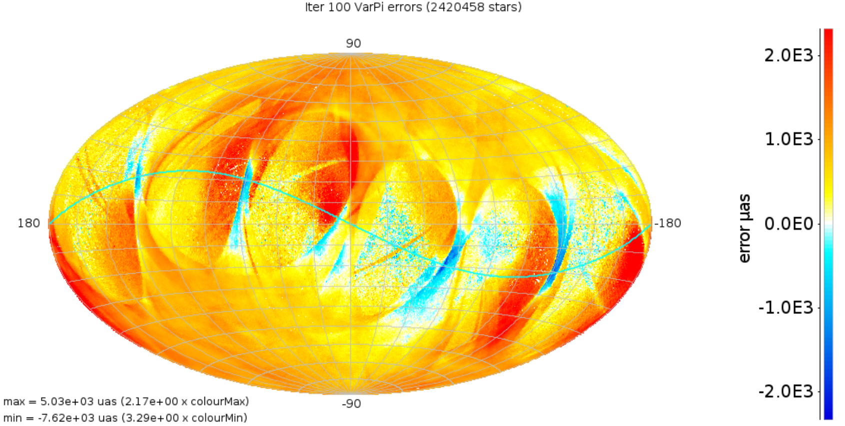


- Basic angle (BA), $\Gamma = 106.5\,\mathrm{deg}$
- Stability critical for absolute parallaxes
- Gaia has on-board metrology, the Basic Angle Monitor (BAM)
- **BAM data shows large variations (approx. 1 mas)**

Figure: L. Lindegren
Publicly available from A. Mora et al. (2014, SPIE)

Results with BA variations

- Simulating variations found by BAM
- Without corrections ⇒ large systematics in parallaxes



Parallax errors (average per pixel, overall median ∼ 0.8 mas)

Figure: D. Michalik. BAV implementation by L. Lindegren & A. Bombrun

Verification of TGAS parallaxes through quasars

Are BAM measurements real? (Michalik & Lindegren 2016)

But: Independent quasar solution in TGAS not possible

1. Add prior for quasars: Assuming zero proper motion
2. Compare resulting parallaxes to zero (BAM expectation: 871.9 μas)

Demonstrated in simulations

Subset	Median [μas]
with spurious proper motions	
Stars	**872.0** ± 0.2
Quasars	**877.7** ± 3.4
with 5% contamination	
Stars	**872.0** ± 0.2
Quasars	**872.0** ± 2.4

Table: excerpt from Michalik & Lindegren 2016, Table 1

Step 7

Conclusions

Summary of thesis research

Study of how to handle stars with insufficient observations

- Scenarios: first release, transient sources, sources at detection limit

Generic prior for non-Hipparcos and non-Tycho stars

- Ensures sensible position estimates and uncertainties

Tycho-Gaia: long-baseline astrometry, full five parameter

- Preliminary results (real data!) very exciting and promising
  - 2.2 million parallaxes and proper motions, Hipparcos-like quality
  - 1 million of very high quality ($\sigma_{\varpi} < 0.32$ mas)
  - Independent parallaxes and proper motions, incl. Hipparcos stars
- Long-period exoplanets from $\Delta\mu = \mu_{\text{TGAS}} - \mu_{\text{Hipparcos}}$
- Challenges: scientific validation, basic angle variations
- Quality of parallaxes can be verified through quasars