Gaia
Data validation

Genius Kick-off meeting, UB
December 4-5, 2013

F. Arenou, CNRS/GÉPI, Observatoire de Paris
D. Hestroffer, IMCCE, Observatoire de Paris

Why a validation in CU9?

Background

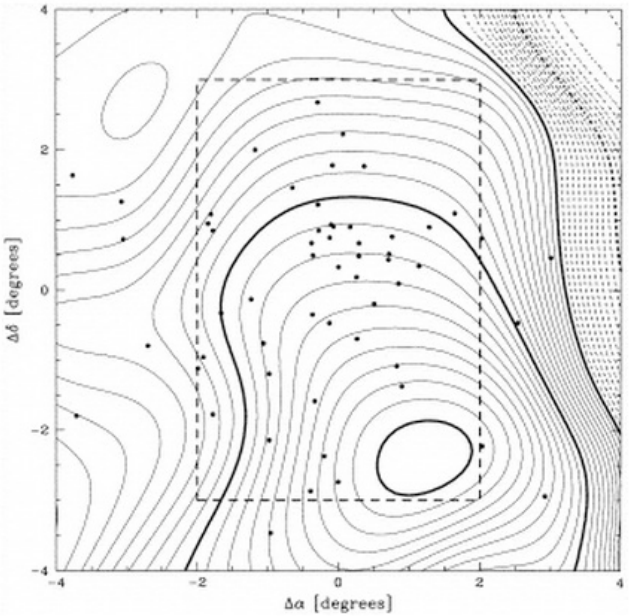
- Gaia is a very complex mission
  - The satellite is a complex engine measuring a complex sky!
  - Obtaining the billions of parameters is a complex process
    - There are many ways to get systematic errors!
- DPAC is responsible of the quality of the Catalogue
  - 450+ scientists/engineers… hundreds of person-years
  - The Gaia Catalogue should not be a quick and dirty work
  - Pressure from outside should not impose the agenda
    - Some form of validation before publication is needed!
- Experience from Hipparcos
  - Users easily misinterpret the (statistical by nature) data
  - Some effort was put in data validation (1PhD, 2 papers, 3 chapters)

THE ASTRONOMICAL JOURNAL, 129:1616–1624, 2005 March

CONFIRMATION OF ERRORS IN *HIPPARCOS* PARALLAXES FROM *HUBBLE SPACE TELESCOPE* FINE GUIDANCE SENSOR ASTROMETRY OF THE PLEIADES[1]

DAVID R. SODERBLOM AND ED NELAN
Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218; sode

G. FRITZ BENEDICT, BARBARA MCARTHUR, IVAN RAMIREZ, AND
McDonald Observatory, University of Texas, Austin, TX 78712; fritz@a
mca@astro.as.utexas.edu, ivan@astro.as.utexas.edu, spies@astro.a



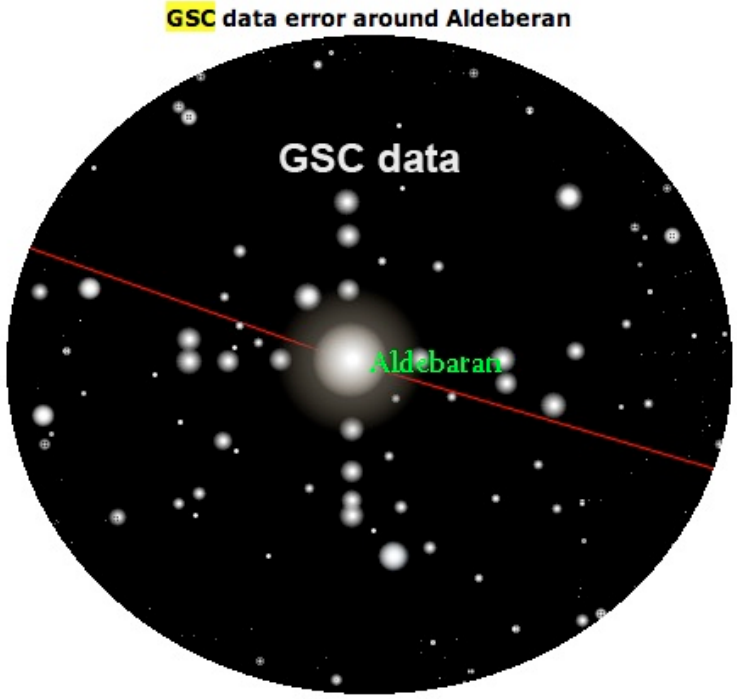
A&A 439, 805–822 (2005)
DOI: 10.1051/0004-6361:20053192

Rights and wrongs of the Hipparcos data

A critical quality assessment of the Hipparcos catalogue

F. van Leeuwen

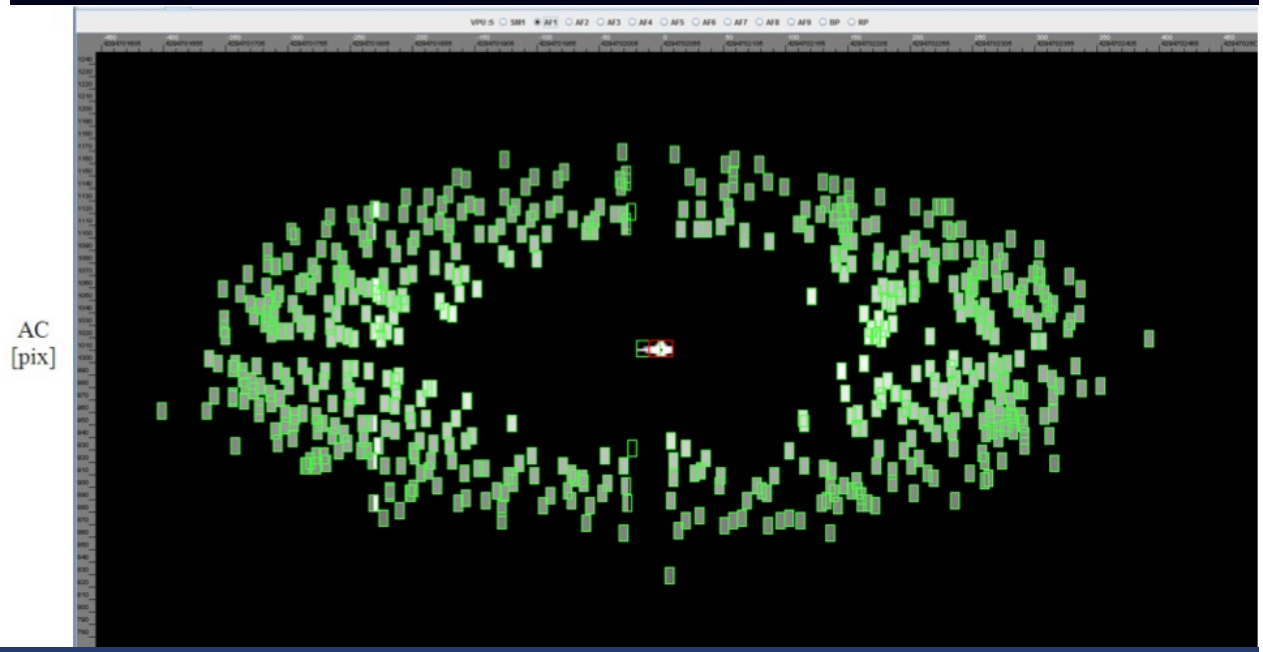
Artefacts…



GSC 1.0 problems

- Many possible problems to expect, e.g.
  - Scanning anomalies
  - Basic angle variations
  - Thermal control anomalies

False detections with Gaia



What validation is

- Cross-CU check of the quality of the Catalogue
  - Have a critical look at the output
  - Do not leave gross errors undetected before publication
  - And correct problems — if any — as soon as possible
    - Feedback to CUs between intermediate Catalogue releases
- Assess the statistical properties
  - Hopefully
    - Unbiased parameters (look for systematic errors)
    - Unbiased parameter standard errors (assess precision, correlations)
    - Find outliers
  - Explain main features in the data
  - Validation results are an integral part of the documentation
    - Documenting the Catalogue properties

What validation is not

- Not a verification of the CUx workflow
  - Already (and much better) done within CUs
  - Though an indirect verification of CU9 tools for the archive access
- Not infallible
  - Minor problems in such a large set of data may remain
- It is not scientific research in our field of expertise
  - But it will require your scientific expertise to understand the tests
  - CU9 is part of DPAC and is bound by the Science Management Plan which states that there will not be proprietary data rights for Gaia.

Validation in other CUs – cont.

- Verification and validation are a common concern
  - Specialized tasks are already on-going
  - Still, more a verification/diagnostic than a validation aspect
- Concerning the publication
  - Results from internal CU verifications should also be documented
  - As part of the description of the Catalogue construction
- Concerning the CU9 validation task
  - The internal CU validation effort should not be duplicated
  - Tools developed within CUs should not be duplicated either
  - All CUs should be represented in the CU9 cross-CU validation work
- There is one (or two) representative from each CU

How to proceed ?

How

- Validation will occur at each intermediate release
  - The complexity of validation being proportional to the scientific content of the release (~ number of parameters and precision)
  - With emphasis of course on the most important data
    - ➢ Our software development process should account for this
- A lot of routine scenarios will have to be implemented
  - Indicating what to test (and why) and what to do when tests fail
  - A VTS (validation test specification) document is being written
- CU9 validation approach should be transversal
  - *Instruments* already handled by Coord Units (astro/photo/spectro)
  - *Objects* sometimes handled by C.U. too (CU4, CU8)
  - CU9 validation will thus mostly be global, based on scientific topics with data being mostly the *combination* of individual C.U. data
    - ➢ We should test what no one else tests…

Run against the clock

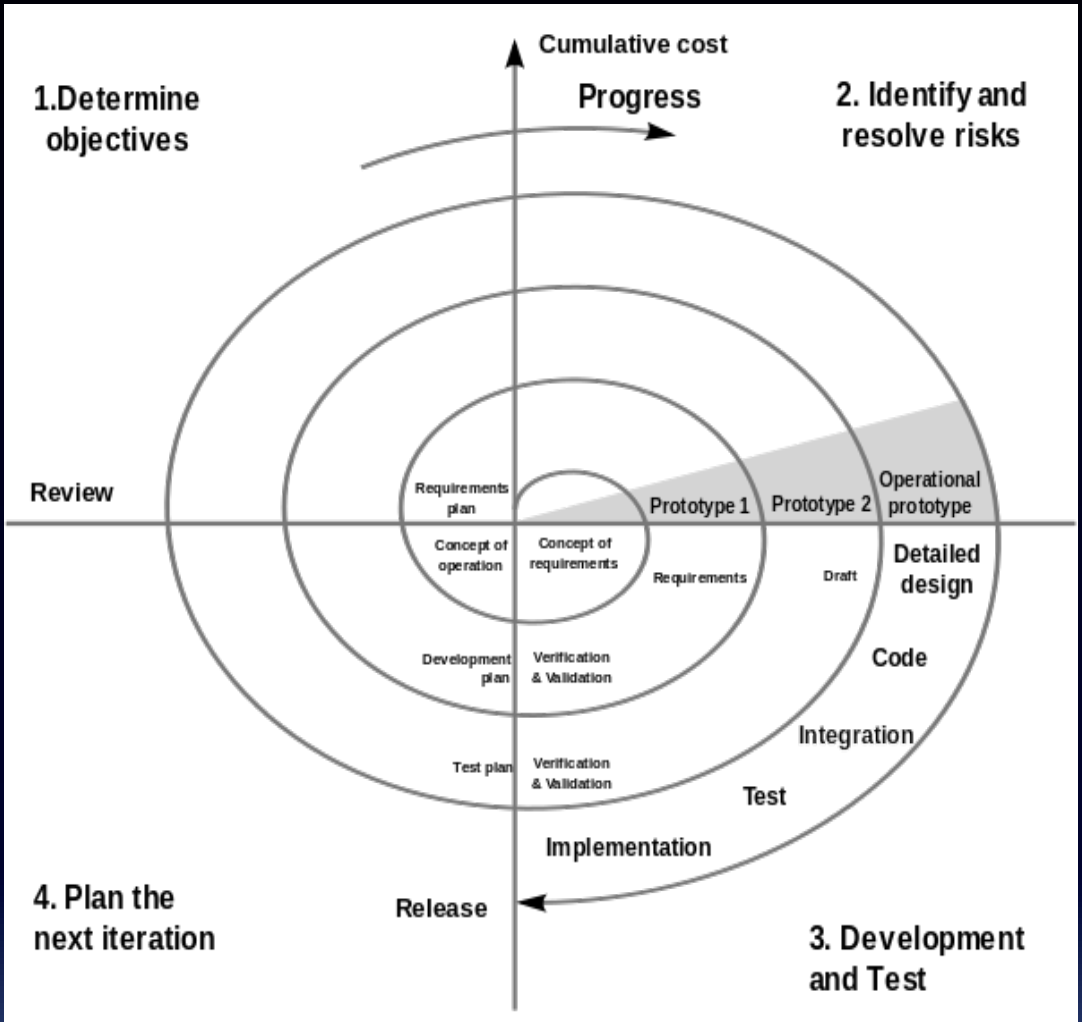
- There will be few time available for validation before each release
- Therefore we must introduce
  - Statistical tests
  - Reporting texts and graphs for diagnostic
    - Use already some notation for the test documents, e.g. 942-12.pdf
- From tests
  - Which should run on routine, with results easy to analyse
    - If no problem found, OK
    - Otherwise, may it come from Gaia/DPAC biases ? Build new tests !
    - If yes, feedback to CUs. Else stop! That is part of exploitation.
  - We may have to `tag' which data should not be published (metadata)
- The tests should themselves be tested through simulations

We will need (simple) blind tests

- Our validation tests are statistical tests
  - How to be certain that the validation tests may indeed detect systematic errors ?
  - What will be the false alarm rate ?
- We will have a simulated catalogue at hand (GOG)
  - Will be used for the type I errors (false positive)
    - The simulations don't contain biases
  - We will use other simulated catalogue, adding biases on purpose for finding the type II errors
    - Checking whether the biases are found by our algorithms
  - Results to be put in the STS/STR

Spiral model

- Each release will see
  - Improvements of the tests which are implemented
  - A better coverage of the tests
- E.g. at first release
  - Spatial distribution uniformity
  - Magnitude distribution as expected in various fields



Practical details

- Software
  - under svn as any other DPAC software (not on local computers!)
  - software language: java as any other DPAC software ?
  - software will be running on routine at ESAC
    - Also guarantees the transparency about what is being done with data
    - Transparency: which tests are done, why & what are the results

- Access to data:
  - TAP interface + (tools for data mining ?)
    - If some data not to be published at some release, then a new release
    - with reason why (traceable documentation)!
  - Also needs to fulfill the needs for early visualisation

Common framework

- Rationale:
  - Most tests will need to be automated and run on routine at ESAC
  - Easier to run the WP94x tests in a consistent manner
  - There will be the need to verify the tests on simulated data (STS)
- What
  - Software and configurations
  - Input, output, results of the tests
- How: what was already done within DPAC
  - SVN, Mantis, LPGL, etc.
  - Java, libraries, etc

- ☑ Action CU9-940TM1-1 (standard file tree)
  - ☐ This has been done, see e.g. http://gaia.esac.esa.int/dpacsvn/DPAC/CU9/software/validation/WP942/

Practical aspects, DB

- Link with CU9 WP930 (archive archi.) = GENIUS WP300
  - I-Chun Shih (OPM) is CU9 WP940 (= GENIUS WP500) interface

- Archive archi. items relevant for validation
  - Validation interface control requirement filled
    - http://www.rssd.esa.int/wikiSI/index.php?title=WP930-interface-control&instance=Gaia
  - Suggestion of CU9 data model
    - There will be an independent CU9 DM, not just an extension of GUMS (CU2). Suggestions/requests from other WP9xx are welcome.
  - Database collaboration
    - Some of us may be doing DB experiments.
    - We do this in Meudon with the Mastodons project (in a small scale)

What validation items ?

Genius vs CU9 Work Packages

- 520 – Looking for trouble
  - F.A./S. Boudreault (CNRS/OP-GEPI)
- 530 – Simulation vs reality
  - A. Robin (CNRS/Utinam)
- 540 - External catalogues
  - C. Babusiaux (CNRS/OP-GEPI)+CSIC+KU
- 550 - Statistical & graphical
  - CNRS/OP-GEPI+FFCUL
- Transversal (special obj.):
  - 563 – Variability
    - L. Eyer (UG)
  - 562 - Multiple stars
    - D. Pourbaix (ULB)
  - 561- Solar system objects
    - D. Hestroffer (CNRS/OP-IMCCE)
- 942 – Scenarios
  - C. Fabricius (UB)
- 943 – Comparison models
  - A. Robin (CNRS/Utinam)
- 944 - External catalogues
  - C. Babusiaux(CNRS/OPM)
- 945 - Statistical & graphical
  - M.Manteiga(OAC)-A.Helmi(Groningen)
- Transversal (special obj.):
  - 946 – Variability
    - L. Eyer (UG)
  - 947 - Clusters as tools
    - A. Vallenari (INAF)
  - 948 - Solar system objects
    - F. Mignard (OCA)

Work Packages

- CNRS except otherwise indicated
- WP520 -Scenarios–
  - Formal validation of the Catalogue field content as function of the object type
  - Internal consistency tests
  - Tests based on what is known to produce effects on given parameters
  - Generation of validation reports with diagnostics filtering
- WP530 - Comparison with models –
  - Extracting the expected statistical properties of the Catalogue from the models
  - Comparing these distributions with the output Catalogue.
  - Retrieving from data already known specific structures;
  - Developing tests for special objects.
- WP540 –Comparison to external catalogues–
  - X-matching, VO ← CSIC, 541
  - Photometric transformations
  - Special areas, special objects
  - Cross-validation tools with Nano-JASMINE data ← KU, 543

Work Packages

- WP550 - Statistical & graphical analysis –
  - Derivation of diagnostics from graphics ← FFCUL
  - Tools using truncated, censored or correlated data, and robust to outliers
- WP561 – Multiple stars – ← ULB
  - Statistical behavior of the solutions leading to the catalogue, standalone
  - Validations based on a comparison with some auxiliary data.
- WP562 - Solar system objects –
  - Confusion between stars and asteroids, perturbations
  - Cross-check of astrometry, RP-BP spectra, photometry
  - Global dynamics of the Solar System, backward computation of occultations
  - Global spectral properties: Gaia taxonomy, against current taxonomy
- WP563 - Variability – ← UG
  - determine if variability behavior is due to the instrument or the reduction,
  - periodicities linked to the sampling law or stars having identical periods.
  - variability trends and correlations with astrophysical parameters

WP520: Internal consistency

- Basic checkings: formal validation
  - Parameter content (check NaN, types, etc.)
  - Subfields present as indicated, e.g.:
    - epoch data present (when and only when indicated)
    - RVS data present as indicated
  - All fields are within valid ranges
  - Check for outliers

- Internal consistency
  - Use assumed properties of parameters (e.g. positivity)
    - No large proper motions for distant stars
  - Exploit intrinsic redundancy between instrument data
    - E.g. photometry should be consistent with spectroscopy
    - Gaia is an complete observatory in orbit!

WP520: Problem-based tests

- Build scenarios based on what is known to produce errors on given parameters
  - Instrumental or calibration problems
  - Classification errors
  - Processing shortcuts, rough models
- Examples, to be more specific
  - Analysis of the variability properties both spatially and in time
    - as photometric calibration problems introduce a spurious variability
  - Check the distribution of parallaxes
    - Annual thermal or calibration effects would introduce a parallax bias
  - Compute distributions of distance to nearest neighbour
    - Components only (possibly redundancies?)
    - Components + sources (possibly redundancies?)
    - From SSO observations to nearest non-SSO (redundancies?)

WP530: Model-based tests

- Develop code on Gaia *simulated* data
  - Extract expected "properties" for all observables
    - Distribution, confidence intervals, ranges for all parameters
    - Correlations between these observables
  - Understand and explain the main structures (see e.g. Hipp. Vol 1)
- Apply this code on actual Gaia catalogue data
  - Apply statistical tests
  - Checking whether the large, expected structures are present
  - Not going into details

WP530: Model-based tests

- A large work already done !
  - Simulations by CU2 will be handy
- CU2 output
  - Universe model
    - Based on Besançon Galaxy model
    - With large add-ons (variable, binaries)
  - Gaia Analysis Tool - GUMS
    - Produces statistics (numbers) or tables to which data can be compared…
  - Add to this: specific models
    - E.g. for solar system objects

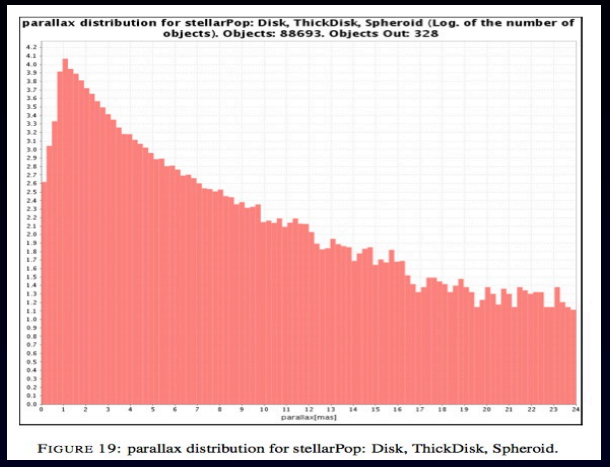


FIGURE 19: parallax distribution for stellarPop: Disk, ThickDisk, Spheroid.

Parallaxes (top)
[Fe|H] vs velocity (bottom)

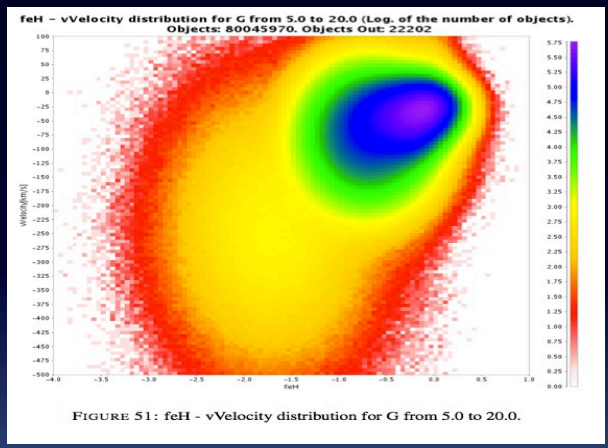


FIGURE 51: feH - vVelocity distribution for G from 5.0 to 20.0.

WP540: External tests

- A very simple recipe
  - Get external data
  - Make cross-matching
  - Compare to Gaia data

- Slightly more complicated in practice !
  - Difficulties to find equivalent data
    - E.g. for astrometry, lack of precision, high level of systematics
    - One reason why Gaia will be launched!
  - Difficulties to X-match
    - No other all-sky survey with a comparable angular resolution and similar multiple star discovering power
  - Difficulties to compare
    - Should not attribute to Gaia, errors coming from comparison data!

WP550: Statistics & Visualisation

- Tests will be statistical
  - Blind tests : e.g. testing systematically ranges of observables
- An effort of fast visualisation is needed
  - All CU2 GAT graphs
  - By epoch or temporal variations
- Comparisons will be far from obvious
  - Beyond scientific competence, statistical analysis skills are needed
  - E.g. working with truncated, censored or correlated data
    - Limited magnitude range, relative precision censorship

Typical validation scenarios (not exhaustive)

Basic checks (examples)

- Subfields present as indicated, e.g.:
  - epoch data present (when and only when indicated)
  - RVS data present as indicated
- Distributions of distance to nearest neighbour, e.g.:
  - components only (possibly redundancies?)
  - components + sources (possibly redundancies?)
  - from SSO observations to nearest non-SSO (redundancies?)
- Fields
  - all fields are within valid ranges
  - all fields have "reasonable" distributions
  - check for outliers
  - for some fields checks may have to be made separately for different classes of sources

Global checks (examples)

- Sky distributions, e.g.:
  - all sources, except components
  - sources with G<20^m, except components
  - median errors for various quantities for various groups of sources
  - distributions of significantly negative parallaxes
- Characterisation of the bright limit
  - which bright stars are missing
  - check surroundings of bright sources for artifacts
- Characterisation of the faint limit
  - will depend e.g. on the number of transits
- Proper motions
  - High proper motion stars are successfully recovered
  - Proper motions for sources with very small parallaxes

Parallax comparisons

- What has been done two decades ago for Hipparcos++
  - Mostly based on positivity
  - Existing ground-based data otherwise very poor
  - Photometric parallaxes + statistical ML model (truncated data)
  - Distant stars

- From that we get a confidence in the data (on a global scale)
  - Parallax systematics + standard errors correctly estimated
    - Now the correlation at small angular scales will be more scrutinized!
  - Need for systematics < 0.1 μas
    - Because data will be averaged, hoping to improve with $1/\sqrt{N}$
  - Checking systematics at the 0.1 μas level yet difficult to achieve
    - Need 5000 bright stars… or 10 million 20^m stars (σ=0.3 mas/star)
    - Using all detected quasars < 20^m I expect a 0.4 μas level only

Typical comparison data used

- Stellar kinematics
  - Which contains both astrometric and spectroscopic data
- Rough consistency for main galactic populations
  - Between position / kinematics / chemical composition
- HR diagram for special populations
  - mixing astrometry + photometry
- Cepheids and other distance indicators
  - Astrometry+photometry +variability

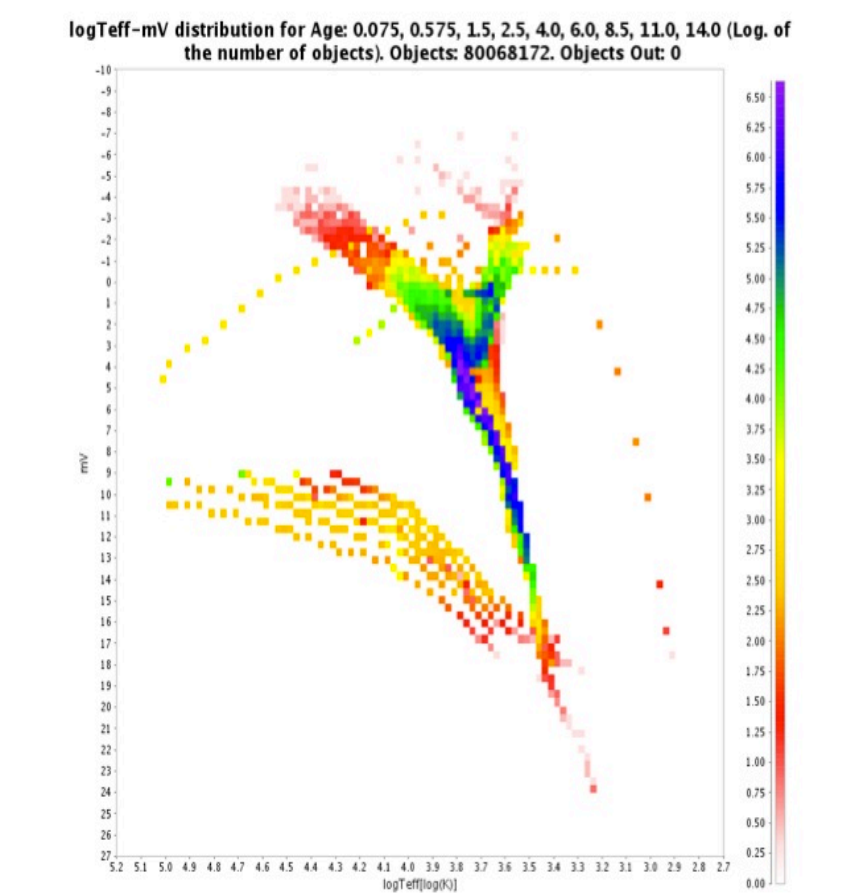


FIGURE 12: logTeff-mV distribution for Age: 0.075, 0.575, 1.5, 2.5, 4.0, 6.0, 8.5, 11.0, 14.0.

Spatial tests (e.g.)

- Production of 3D spatial maps
- Analysis of the on-sky (2D) spatial distribution of poorly classified objects or non-classified objects
  - e.g. low DSC probability
  - and their spatial neighbourhoods,
  - to see whether photometric delending/crowding problems may be an issue (this could feed back into improving BP/RP extraction)
- Analysis of the 3D interstellar extinction distribution
  - Compared to our current understanding of gas and dust distributions from infrared surveys (colours)
- Analysis of the Galactic metallicity distribution
  - both spatially and as a function of stellar kinematics and ages (produced by CU8)

Luminosity tests (example)

- H-R diagrams for selected stellar populations
  - e.g. known globular/open clusters, compared to current knowledge
- G-band absolute magnitude function
  - perhaps the luminosity function too, with the APs calculated in CU8
  - For various samples of stars + compare with current knowledge.
- QSO redshift and luminosity distribution
  - Compared to results from SDSS, Pan-STARRS and other surveys,
  - Taking into account the selection effects.
  - This will help understand the type I and type II errors in the CU8 QSO/star classification

Organisation

Within CU9/GENIUS and outside

- The various CU9/GENIUS work areas are not independent:
  - E.g., the validation will need the tools developed within CU9
  - These tools depend on the Operations and Support area.
- Validation will use analysis tools developed within GENIUS
  - Can also indirectly be a validation of the analysis tools.
  - This has to be accounted for in the work package definition and in the timescales (e.g. validation will need some own tools before)
- Also an interaction/feedback with other CUs is needed
  - Not only CUs: the "Science Alert" work is also a validation task
    - ➢ False positive may be due to calibrations, etc. problems

CU9 Schedule until first release

- First Cycle (A):
  - 10m from now till March 2014
    - Start up, SDP, SRSs, ICDs: Before the end of 2013
- Second cycle (B):
  - 5m long, ending early Sep 2014.
  - End of Sep : first public release of GACS (1.0) and MDB-00 will be available for validation
- Third cycle (C):
  - 5m long ending end Jan 2015.
  - 6m for validation before first release@L+22m: 07/2015
  - MDB-01 release at L+19 months, April 2015. (JSH-033)
  - 3m for final validation!!

Releases

- I First release: Launch + 22 months (July 16, 2015)
  - 90% Positions and G magnitudes for single stars with error estimates accounting for calibration errors
  - 100 000 Proper Motions (HTPM) catalogue based on the Hipparcos stars
- II Second release: L + 28 months (> start 2016)
  - Positions (+parallaxes+pm where available) and G magnitudes for single stars with good errors
  - Integrated photometry BP/RP with verified basic astrophysical parameter estimation
  - 90% Mean radial velocities for constant stars
- III Third release: L + 40 months (> start 2017)
  - 90% Five parameter astrometric solution for single stars
  - Binaries: orbital solutions 2 month < P < 75% observation
  - RVS spectra and Spectrophotometry from BP/RP for sources with astrophysical parameters
  - Source classifications (probabilities) plus stellar effective temperatures and extinction
- IV Fourth Release: L + 65 months (> start 2019)
  - Updates of astrometry with appropriate error estimates.
  - Spectrophotometry from BP/RP for sources for which astrophysical parameters are released.
  - Mean RVS spectra for sources where single epoch spectra are usable.
  - Source classifications, stellar astrophysical parameters for the majority of stars.
  - Orbitals solution for periods between 2 months and 75% of the observation duration
  - Variable star classifications and parameters as available, epoch photometry .
  - Solar system results with preliminary orbital solutions and individual epoch observations
  - Non-single star catalogue

CU9 Milestones

- Tests will be described in a Validation Test Specification doc.
  - On-going

- CU9 Deliverables
  - Definition of tests
  - Software code ← GENIUS
  - Results from tests
  - Reports

- The tests define the CU9 milestones, 2/3 milestones per test
  1. The intermediate steps to achieve the algorithms
  2. The SW code released ← GENIUS
  3. The tests completed and documented

First CU9 Milestones

WP942.1	Cycle A	Check of parameter ranges, NaN, etc.	30-mars-2014
WP942.3	Cycle A	Effect of contaminations (bright stars, etc.)	30-mars-2014
WP942.5	Cycle A	Subfield or other data present when indicated	30-mars-2014
WP943.1	Cycle A	Spatial distribution of magnitudes vs galactic model	30-mars-2014
WP943.3		MPA distribution of proper motions vs galactic model	
WP944.1	Cycle A	Spatial distribution of magnitudes vs external catalogues	30-mars-2014
WP944.3	Cycle A	MPA distribution of PM vs external catalogues	30-mars-2014
WP945.1	Cycle B	Outliers search (HTPM)	30-sept.-2014
WP947.1	Cycle A	test design and external catalogue definition	30-mars-2014
WP947.3	Cycle B	mag + pm on Hipparcos data	30-sept.-2014
WP948.1		backward computation of approaches etc. TBD	

GENIUS deliverable

Del. Num.	Description	Month
500.1	Delivery of prototype of internal checking tools (WP 520)	12
500.2	Delivery of prototype of statistical tools (WP 550)	18
500.3	Delivery of internal consistency checking tools (WP 520)	24
500.4	Delivery of statistical tools (WP 550)	24
500.5	Delivery of model-based validation tools (WP 530)	24
500.6	Delivery of prototype of external validation tools (WP 540)	24
500.7	Delivery of external validation tools (WP 540)	36
500.8	Delivery of special object tools (WP 560)	36
500.9	Deployment of validation tools on the Gaia archive	42

CNRS

- The French National Centre for Scientific Research
  - the largest governmental research organisation in France
  - the largest fundamental science agency in Europe.
  - It employs 26,000 permanent employees (researchers, engineers, and administrative staff) and 6,000 temporary workers.
- For Genius, CNRS is an umbrella for
  - Observatoire de Paris/GEPI (UMR 8111)
  - Observatoire de Paris/IMCCE (UMR 8028)
  - Observatoire de Besançon/Utinam (UMR 6213)
- Through its Mixed Research Units (UMR)
  - Mixed funding, mixed staff

Observatoire de Paris

- PARIS : founded in 1667 by Louis XIV & Colbert
  - Building by Claude Perrault
- MEUDON : observatory dedicated to the physical astronomy
  - Founded by Jules Janssen in 1876
  - Attached to Observatoire de Paris in 1926
- NANCAY : dedicated to radioastronomy
  - Founded in 1953 by ENS researchers

- Five Departments,
  - more than seven hundred people
  - over the whole research spectrum in Astronomy & Astrophysics
- One of the largest Institute

CNRS/Obs. Paris-GEPI (Meudon)

- One of the Department of Paris Obs… in Meudon.
- The main research topics : the formation and evolution of stars in our Galaxy as well as in numerous other galaxies
- This research calls upon many disciplines, from chemistry to physics, from instrumentation to data-processing engineering, and from project management to financial management within an international frame-work.
- Frédéric Arenou will lead WP500, Paola Di Matteo is an expert in numerical simulations and will be a key person in WP 500, as well as Carine Babusiaux, deputy manager of the DPAC data simulation unit CU2.

CNRS/Obs.Paris-IMCCE (Paris)

- An institute of Paris Observatory.
- IMCCE is also under the umbrella of the French Academy named Bureau des longitudes.
- IMCCE researchers focus on studies of the Solar System and planetary systems in the domains of celestial mechanics, astrometry, planetology and mathematics.
- Also, IMCCE is in charge of providing the national ephemerides in France.
- Jérôme Berthier, team leader of a DPAC CU4 development unit, is an expert in the dynamics of asteroid systems and VO aspects and will contribute to this latter aspect in WP300 and WP400, while Daniel Hestroffer, director of the IMCCE, will manage the tools specialised in Solar System Objects validations in WP500 together with William Thuillot, former IMCCE director and head of the Gaia-FUN-SSO ground-based follow-up network.

CNRS/Utinam (Besançon)

- UTINAM = Univers, Transport, Interfaces, Nanostructures, Atmosphère et environnement, Molécules
- Unité Mixte de Recherches du CNRS (UMR 6213) et de l'Université de Franche-Comté
- The Institute studies the structure and dynamics of isolated systems and interactions with complex environments, over a wide range of space and time scales. The large scale systems studies include our galaxy, the Solar System and single, and double astronomical objects.
- An important associated activity is the establishment and dissemination of accurate standards for the measurement of time and frequency, as well as the discovery of useful extragalactic astrometric references for spatial measurements.
- Annie Robin and Céline Reylé, leaders of the DPAC CU2 simulated universe model, are the most competent to take charge of the aspects of the Gaia data validation using a Galaxy model.

CNRS involvement background

- The CNRS centres participating in GENIUS have been involved
  - from the very beginning in the definition of the Gaia mission,
  - the simulation of its scientific content and the data processing within DPAC.
- In the past, the CNRS group at the Paris observatory was in charge
  - Of one of the Hipparcos Consortia
  - of the validation of the Hipparcos catalogue, and therefore brings an invaluable experience in this field for WP500 (catalogue validation).

CNRS involvement in Genius

- WP-330: Deployment of specific web services
  - the SkyBOT12 service suite + Miriade13 ephemerides
  - 1.8 staff month
- WP-450: Comm. portal, outreach academic
  - Journey through the Galaxy, exhibition material
  - 1.8 staff month

- WP-500: Tools for data validation and analysis
  - Management+Tools
  - 75.6 person month
    - CNRS/IMCCE: J. Berthier, D. Hestroffer, W. Thuillot + 2yr hired
    - CNRS/GEPI: F.Arenou,C.Babusiaux,P.DiMatteo, I. Shih+ 3yr hired
    - CNRS/Utinam: C.Reylé, A. Robin+ 1yr hired

Funding/CNRS total

	RTD	DEMO	TR	MGT	OTH	Total
Personnel	339 881			0	14 612	354 494
Subcontracting	0					0
Other direct costs	15 000			0	0	15 000
Indirect Costs	212 929			0	8 767	221 696
Total costs	**567 810**			**0**	**23 380**	**591 190**
Requested EC contribution	**425 858**			**0**	**23 380**	**449 237**
Person.months	79				2	81

- Funding mostly for recruitment
  - 79MM+2MM management

Funding/each CNRS institute

	RTD	OTH	Total
Personnel	157 829	14 612	172 441
Subcontracting			0
Other direct costs	5 000		5 000
Indirect Costs	97 697	8 767	106 465
Total costs	**260 526**	**23 380**	**172 441**
Requested EC contribution	**195 394**	**23 380**	**218 774**
Person.months	37	2	39

GEPI

	RTD	OTH	Total
Personnel	126 240		126 240
Subcontracting			0
Other direct costs	5 000		5 000
Indirect Costs	78 744		78 744
Total costs	**209 984**		**209 984**
Requested EC contribution	**157 488**		**157 488**
Person.months	28		28

IMCCE

	RTD	OTH	Total
Personnel	52 355		52 355
Subcontracting			0
Other direct costs	5 000		5 000
Indirect Costs	34 413		34 413
Total costs	**91 768**		**91 768**
Requested EC contribution	**68 826**		**68 826**
Person.months	13		13

	RTD	OTH	Total
Personnel	3 458		3 458
Subcontracting			0
Other direct costs			0
Indirect Costs	2 075		2 075
Total costs	**5 533**		**5 533**
Requested EC contribution	**4 149**		**4 149**
Person.months	1		1

Utinam+3rd party

CNRS organisation+funding

- CNRS will be the main provider of WP 500,
  - 5% together on the management and administration of this Work Package for a total of about 2 months.
  - a part of the CNRS involvement will also take part in the VO tasks of WP 300 and WP 400.
- Staff positions + a postdoctoral level on one hand and two engineer levels for implementation of the validation tools.
- Travel funds for the meetings within the CNRS partners or foreign (ULB, UG) nodes working in WP 500.
- Regarding small equipment (e.g. laptops) and consumables, as most CNRS positions are permanent positions for which CNRS already cover the expenses, the requested funding is needed only for hired people.

Current situation

- The good
  - OP/GEPI has already recruited
    - Steve Boudreault
    - Gaia background (MSSL)
  - CNRS has all the needed administrative resources
    - Already one K-O meeting
- The bad
  - Besançon is far from everywhere
    - And too cold, maybe ;-)
    - Difficulty to recruit
    - 2 candidates already resigned


Thank you for
your attention