Taipan Swinburne Meeting, Nov 30th - Dec 1st, 2015

Participants:

- Elisabete da Cunha
- Chris Blake
- Christina Magoulas
- Fuyan Bian
- Kyler Kuehn
- Tamara Davis (Dec 1st)
- Michael Brown
- Chiara Tonini
- Jeremy Mould (Nov 30th)
- Matthew Colless
- Cullan Howlett
- Caitlin Adams
- Andrew Johnson
- Andrew Hopkins
- Ned Taylor
- Sam Hinton
- Nick Tothill (Dec 1st)
- Ixandra Achitouv
- Jennifer Piscionere
- Heath Jones (remote)
- Matt Owers (remote)
- Yjan Gordon (remote)
- Elaine Hyde (remote)
- Aaron Robotham (remote)
- Angel Lopez-Sanchez (remote)
- Mike Hudson (remote)
- Fred Watson (remote)
- Nuria Lorente (remote)

Main outcomes from the previous workshop at the AAO in August:

1. We decided to base the first year (Taipan Y1) on already available photometric catalogues, since it is too risky to wait for SkyMapper. We decided on a two-tier approach for Y1:
   - all-sky shallow based on 2MASS extended source catalogue (extended sources are galaxies so that helps us with star/galaxy separation), also maybe point source catalogue if we can reliably select galaxies. With this we can already have a significant H0 result (need forecasts to demonstrate that), we can re-do the whole 6dF analysis and potentially already have a significant improvement over 6dF (maybe all of the peculiar velocity science can actually be done in Y1; TBC).
   - go to full depth/completeness in a small area: this would be a pilot sample for galaxy evolution science. The field where to do this was not decided on, but KiDS seems like a good option; pending on checking available photometric catalogues etc.

2. Tiling code: we realised that we need a custom-built tiling code for Taipan. Matthew hired Marc White at ANU for a month to work on a tiling code according to a set of specifications we provided him, and this was extremely successful. We now have a working python tiling code.

3. We started shaping up the plans for Science Verification phase, which will start immediately after the instrument commissioning is completed (mid-2016). We decided to observe two fields in this phase:

one ‘known field’ where spectra already exist to perform tests on how spectra (S/N, calibration, etc.); the suggested field for this was Stripe 82. The second field would be a ‘new science field’, where obtaining Taipan spectra can already provide a significant result (good for survey advertising); the suggested field is the SPT Deep Field, where we can obtain optical spectra (redshifts, AGN diagnostics) of radio sources. More work needed on specifying input catalogues and detailed plans.

Goals for this meeting:

1. To make decisions and establish plans for **Taipan Y1**. Specifically, decide on input catalogues (for science targets, guide and calibration stars), work on tiling Y1 to estimate the survey duration and efficiency, work on cosmology forecasts for Y1, to understand what precision on H0 we can achieve after the first year of survey, and address the question of how far we can get with the peculiar velocity science in the first year.

2. To make more detailed plans for **Science Verification**, including a roadmap for all the data that need to be taken and tests that we need to perform. At the last meeting we decided on Stripe 82 ad the ‘known field’, but we needed to decide if we are using the SPT Deep Field as the ‘new science field’. Specifically, are there available photometric catalogues to use as input? Also, discuss the new science (obtaining redshifts for radio sources) in more detail.

3. To start planning **operations and data flow software**. We now have an operational tiling code, but we still need a scheduler, a data reduction pipeline, spectral analysis software, and software to ‘glue’ this all together into a coherent schema. We also need to design and establish the database structure for our data. As discussed before, we probably do not need to re-invent the wheel, but discussion is needed to understand if it is more appropriate for our survey to make any component(s) from scratch (as the tiling code showed).

Matthew — general update of Taipan status:

- refurbishment of the UK Schmidt Telescope finished
- successful ARC DP by Matthew, Andrew and Chris Blake, will cover half of the operating costs of the survey, and pay for two post-docs to work on Taipan (ANU and Swinburne)

Kyler — instrument & software update:

- the positioning of the bugs is going well, precision down to fractions of a micron, exceeding the requirements for guiding
- note: the plate needs to be cleaned because if it is dirty that affects the movement/positioning of the bugs. It will probably need to be done once a week (still TBC, but definitely will not be needed multiple times a night, which is good)
- team is using trello to tack the workflow of testing each individual starbug. Very good tool to manage tasks and team work, may be worth using for Taipan survey planning
- Nuria has set up a database of starbug ‘personalities’: the bugs do not go exactly where you tell them to go, but at least the deviations (i.e. personalities) are consistent and repeatable. So we can use this database to correct the positioning accordingly.
- commissioning of the instrument still ok for mid-2016. Science verification can start right after that (most likely July or August)
- Matthew asked about the definition of the instrument control interface. Not sure yet, we need to ask Keith or Nuria. It would be helpful to get the planned requirements so that we can build the scheduler/ pipeline accordingly.

Ned — recent progress on Y1 and SV photometric catalogues:

- The “input catalogues” working group led by Ned is working on target catalogues for each of the three science goals: cosmology, peculiar velocities, and galaxy evolution. For the cosmology and peculiar velocity goals, volume is more important than completeness, while for galaxy evolution completeness is more important than volume, and value comes also from ancillary data at other wavelengths.
- Discussions and catalogues available on Slack. Ask Ned about it if you would like to join.
- For Y1, Ned investigated the 2MASS (XSC+PSC) all-sky catalogue for the all-sky survey (emphasis on volume) and KiDS (VST Atlas selected; emphasis on completeness) for the smaller area for the galaxy evolution science. After Y1, the expectation is to follow SkyMapper over the Southern hemisphere.
- Ned used near-IR colour criteria to separate stars from galaxies and then used SDSS to evaluate how well this separates stars from galaxies (assuming SDSS separation is perfect). He gets 99.93% success rate.
- The near-IR based selection seems to be doing a great job, but to get more galaxies at higher redshift we will need additional colour selections (?)
- Standard star catalogue (for spectro-photometric calibration) will come from SkyMapper. If SkyMapper is not ready on time, we can use PanSTARRS as plan B
- Summary: we have working catalogues for sample selection and survey forecasting, and star/galaxy separation seems to work well. Next: work on tiling/scheduling strategy.

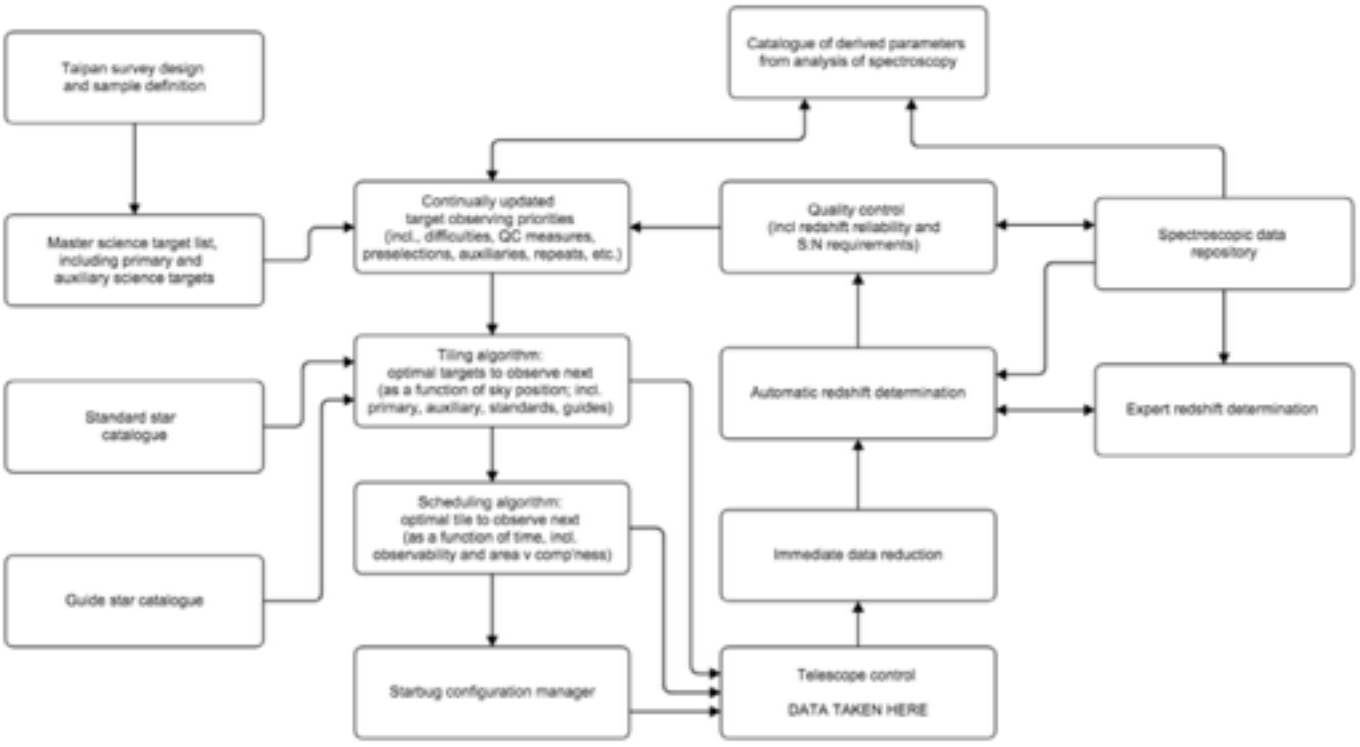


Fig.1: Ned’s data flow chart. This is a nice summary of all the components that need to be in place and how they should talk to each other (but some details still TBC). Credit: Ned Taylor.

Chris Blake — Summary of the Cosmology science case and Taipan Y1 forecasts:

- In brief, the Taipan cosmology science goals are:
  1. to obtain complete cosmological information in the low-z Universe, where dark energy dominates, to complement current and future high-z galaxy surveys;
  2. to make a 1% (model-independent) measurement of the expansion parameter H_0 using the baryon acoustic peak, enabling fundamental tests of the cosmological model;
  3. to perform new tests of GR across a range of scales using two complementary methods, galaxy peculiar velocities and redshift-space distortions.
- **[action item:]** alter the plot that compares Taipan H0 measurements with others; e.g. plot the Planck error bar without assuming a model to really show the impact and improvement of Taipan constraints.
- The Taipan Y1 cosmology survey will be a self-contained 1-year survey of the whole sky (with targets selected from 2MASS) that will target ~400,000 sources.
- Chris made some forecasts to try to understand: 1. what is the optimal selection for this Y1 sample, and 2. if we can produce a 2% measurement of H_0 in Taipan Y1.
- Optimization of the Taipan Y1 cosmology survey (with Ned):
  - maximize the survey volume spanned by ~400,000 targets, given observational limits
  - do not re-observe 6dFGS redshifts in Y1
  - reach higher redshifts by: 1. including 2MASS point sources, and 2. using a minimum J-K cut
  - we do not know optical magnitudes, but we estimate them using Ned's proxy: r = J+1.1+0.8(J-K)
  - the SDSS-matched "sandbox" catalogue gives us redshift distributions and fibre magnitudes
  - use Fisher matrix techniques to predict H_0 and growth measurements given survey N(z) and area
- Below are some interesting plots resulting from the analysis of this "sandbox" catalogue.

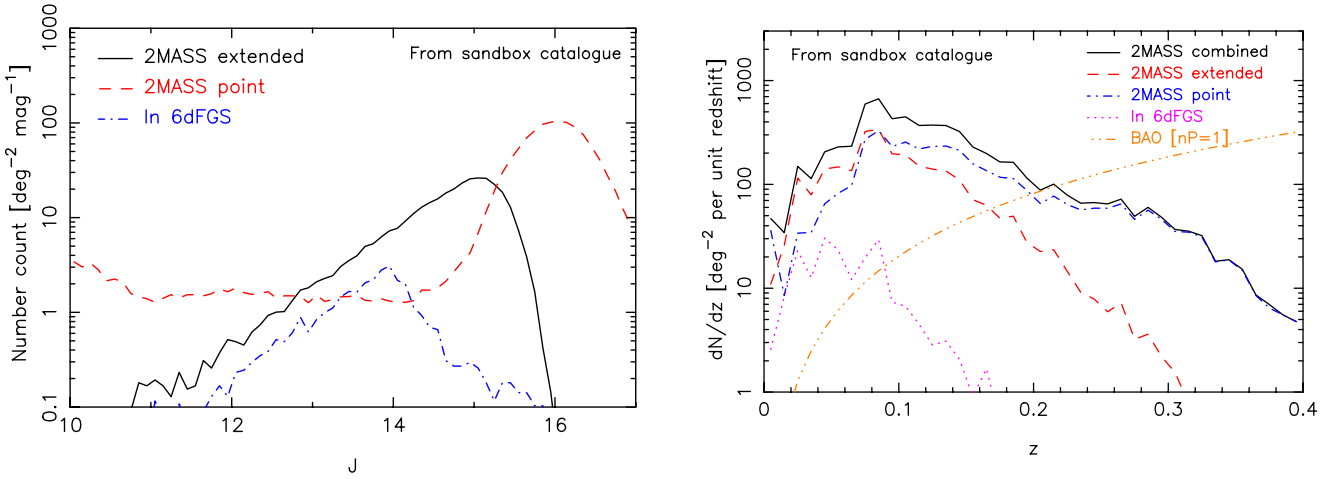


Fig.2: effect of including 2MASS point sources on the J-band number counts (left) and redshift distribution (right). If we include the point sources we can go to higher redshifts and this probe a larger volume. Credit: Chris Blake.

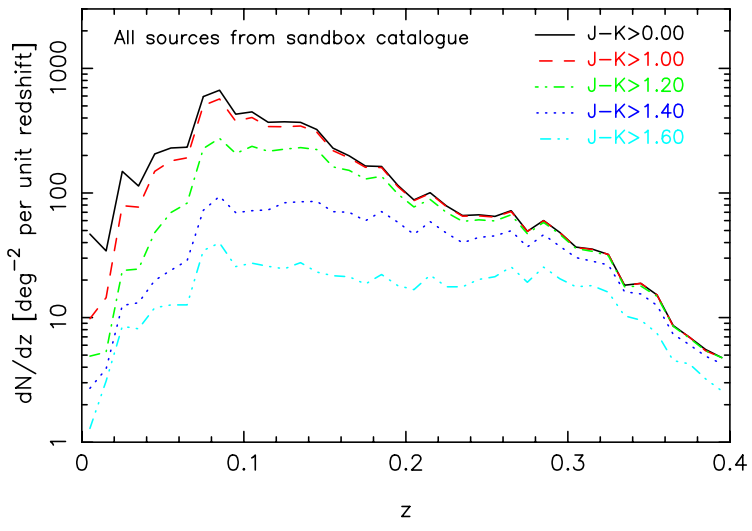


Fig.3: using a J-K colour cut can help us obtain a flatter redshift distribution (i.e. get rid of low-redshift sources that do not really improve the measurement). With a J-K>1.6 colour cut we get a flatter N(z), optimal for H0 science. Credit: Chris Blake.

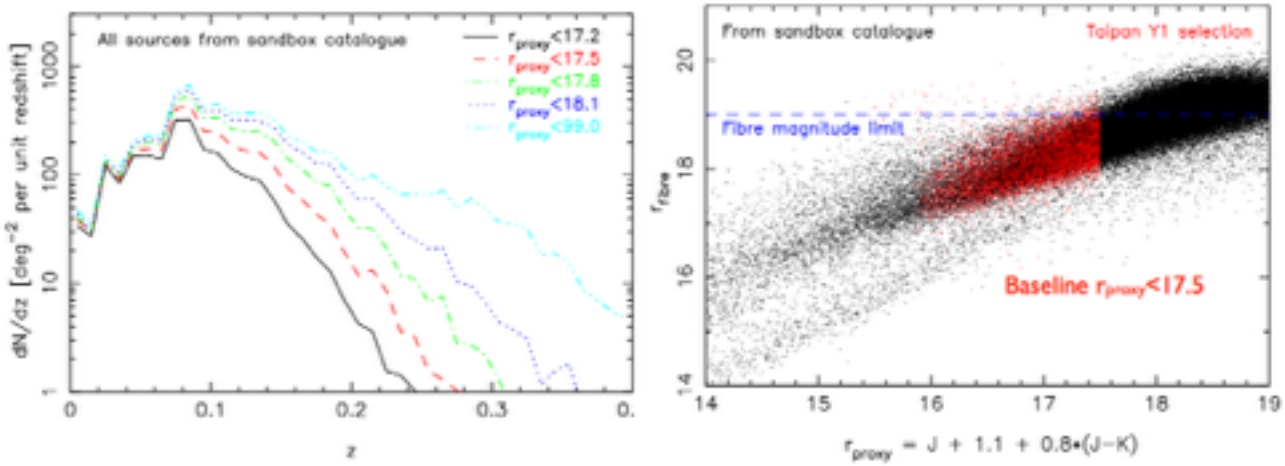


Fig.4: left: redshift distribution as a function of r-band magnitude; right: comparison with fibre magnitudes: this tells us how faint we can really go. Credit: Chris Blake.

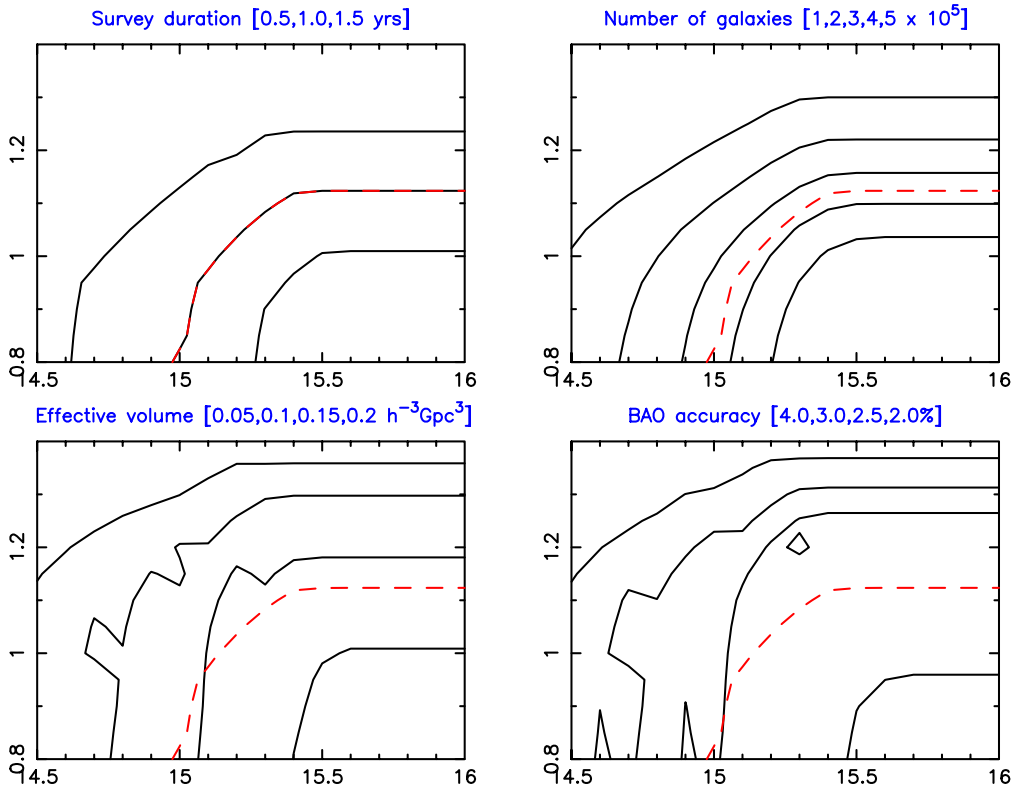
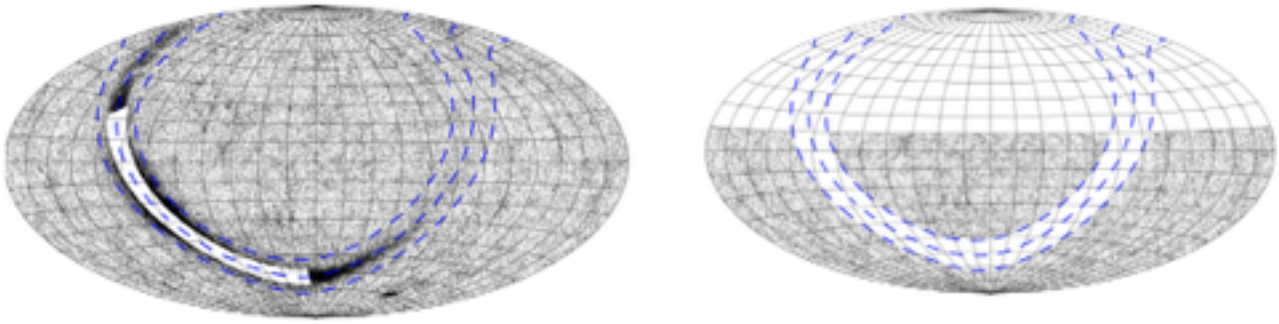


Fig.5: BAO optimization for r_{proxy}<17.5: J<15.5, J-K<1.1 i.e. Taipan Y1 baseline. This is based on a catalogue of 400,000 sources (see Fig. 7), and constrains H_0 to 2% and growth to 8%. Credit: Chris Blake.

J	J-K	r_{proxy}	V_{eff} (Gpc/h)^3	BAO (%)	growth (%)	Notes
<15.5	>1.1	<17.5	0.183	2.1	7.8	Taipan Y1 baseline
<15.0	none	<17.5	0.161	2.3	8.7	no J-K cut
<15.5	>1.0	<17.2	0.146	2.4	8.9	brighter r_{proxy}
<15.7	>1.2	<17.8	0.213	1.9	7.1	fainter r_{proxy}
<15.5	>1.05	<17.5	0.169	2.2	8.1	just extended
<15.5	>1.1	<17.5	0.238	1.9	7.1	adding in 6dFGS

Fig.6: Comparison of different BAO optimizations for Taipan Y1. Credit: Chris Blake.



Cut	Explanation	Number
none	Taipan_InputCat_v0.101_20151125	2,413,252
dec<10, |b|<10	Visibility, avoid Galactic plane	1,020,721
r_{proxy}<17.5	Approximate observational Taipan limit	902,073
J<15.5	NIR limit [not much effect given r_{proxy}]	860,880
K>12.75	Do not re-observe 6dFGS sources	747,376
J-K>1.1	Preferentially restrict to high-z	408,590

Fig.7: Cuts applied to the main 2MASS-based input catalogue to obtain the Taipan Y1 baseline sample. Top: sky distribution of sources before and after cuts. Notes: 15% point sources, 3% fainter than the fibre limit, and unknown stellar contamination. Credit: Chris Blake.

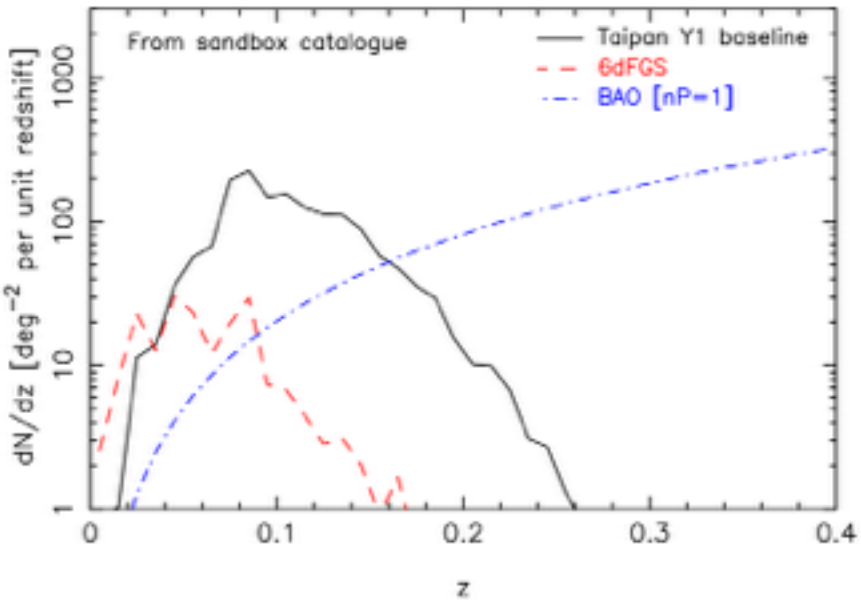


Fig.8: Redshift distribution of Taipan Y1 baseline selection. Credit: Chris Blake.

- Results are model model dependent when going to higher redshifts. **[action item:]** make plot showing where model dependencies start hurting us.
- 90% spectroscopic completeness might be sufficient for Cosmology (efficiency is much more important); the limiting factor reverts to redshift success.
- Next steps:
  - apply tiling code to baseline Taipan Y1 cosmology catalogue to check what is the efficiency of tiling allocation [Ned];
  - to what extent does the Y1 cosmology catalogue allow the Taipan peculiar velocity science to be completed? [Christina, Matthew];
  - generate first version of mock catalogues;
  - apply tiling code to mocks, investigate clustering systematics due to correlation of allocation with density;
  - develop curved-sky BAO reconstruction code;
  - continue cosmological science with 6dFGS.
- Conclusions:
  - Taipan will allow the ultimate low-redshift tests of cosmic expansion and gravity;
  - baryon acoustic peak will measure H0 to 1%, cross-checking CMB vs local standard candles;
  - redshift-space distortions in the galaxy sample will produce the best measurement of the z=0 growth rate, testing GR on intermediate scales;
  - PV sample will allow new tests of GR on the largest scales of 100s of Mpc/h;
  - Taipan Y1 will make major progress in these goals [very encouraging results!]

Ned — Tiling Taipan:

- Marc White's code is working beautifully, the only issue for now is that the code cannot tile the whole sky as it is; for now, Ned broke the sky up into a number of cells that have roughly equal volume to make the tiling manageable;
- Marc's tiling code can be found here: https://github.com/marc-white/taipan-tiling
- no priorities yet, but that is something that can be easily added
- the order by which the tiles are observed needs to be determined — a job for the scheduling code?
- Ned presented a very nice gif showing the tiling progression (both for a cell of the whole-sky catalogue based on 2MASS and for the equatorial strip); this was done using the "greedy" algorithm
- to be most efficient, go for most difficult targets first
- decisions to be made: difficulty vs priority vs completeness (vs observability for scheduling)
- how do we tell the code where to go first? and how do we decide where we want to go first?
- 22% spare fibres
- need for repeat observations in Taipan Y1 not clear yet: needs to be checked in science verification
- would having 150 more fibres help us? still TBD: check by running the code with 300 fibres
- Ned finds that Taipan Y1 is tileable to 95% completeness (modulo scheduling, observability)

Sam — MARZ:

- MARZ is an automated redshift code based on a modified version of the auto-z cross-correlation algorithm by Baldry
- code can be found here: https://samreay.github.io/Marz
- runs in browser, and now also on command line (we can have a command line user interface for Taipan)
- at the moment the code shows only the best redshift template match, but it is very easy for Sam to add multiple matches (how many would be a good number?)

- some templates may be disabled if needed/appropriate (e.g. QSO); this can be tested in Science Verification
- data processing challenge: can we save a single FITS file for each object will all the spectra (and resulting redshifts) for that object?
- Sam showed that thanks to having a QSO template (with broad lines), MARZ does really well with samples that include QSOs
- the comparison slide shows that MARZ is very good at recovering redshifts automatically
- MARZ runs in Javascript to it takes a bit of time (1 sec per spectrum on one core): is this good enough for us or do we need to speed it up? probably not if we run on multiple cores.
- would it be useful to add redshift priors from GAMA (i.e. P(z) as a function of r-band magnitude)?

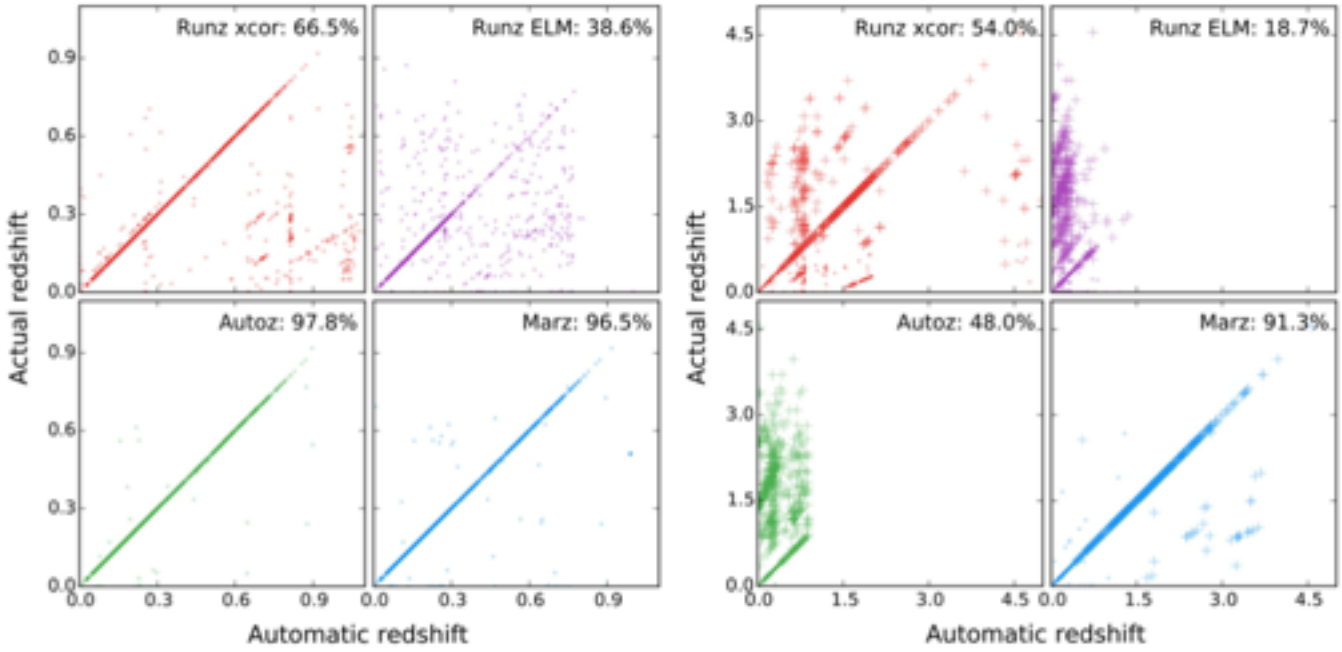


Fig.9: Comparison of MARZ performance with other codes. Left-hand panels include QSO templates, while right-hand panels do not. Credit: Sam Hinton.

Runz	Marz	Autoz
People know how to use it	Good low-z redshifting	Good low-z redshifting
Allows manual redshifting	Good high-z redshifting	Command line running
Bad automatic matching	Easy to install/update	Fast
Hard to update	Allows manual redshifting	Requires IDL
OS dependent	Cmd requires node/iojs	No manual redshifting
Tricky to install	Slow matching	No high-z redshifting

Fig.10: General comparison of MARZ with other commonly used redshift codes. Credit: Sam Hinton.

DISCUSSION: Data pipeline

- data reduction by AAO software team (building on 2dFDR); currently no people working on it; we need someone from Taipan to sit on the 2dFDR working group [Andrew will do it]
- a lot of the work on the data reduction software will be modelling the spectrograph (with instrumentation team)
- we need to write down the requirements for what is needed from 2dFDR effort, as they need detailed specifications
- 2dFDR/Taipan: same kinds of outputs as AAOmega; flat fielding; wavelength calibration; adopt a model for the flux calibration based on Millie's code, in the same format as AAOmega spectra
- **[action item: Andrew]** put together 2dFDR/Taipan specs; join 2dFDR working group
- 2dDFR -> MARZ in a scripted automated way
- processing done within a night is only strictly needed if the calibrations change, otherwise we can do it in the next night (this is crucial for co-adding spectra) [needs to be tested in SV phase!]
- Sam and Matt will co-ordinate output from MARZ to feed as input to MO spectral line measurement tool
- MO input: reads FITS file with info from header: redshift, spectrum, wavelength solution; other versions also go and get the spectrum that 2dFDR outputs; then sends S/N, quality flag back to database
- [SV test:] do the current S/N requirements for each subsample give us the needed precision in redshift and velocity dispersion?
- MO code takes 90 sec/spectrum (PPXF), makes it difficult to process every night
- need mechanism to update target list (database) based on quality flag (q_OP) and S/N
- need an air traffic control system i.e. a database system putting this all together. Sam is very keen to work on implementing this database system
- **[action item: Exec]** put together specs for pipeline and database control system

DISCUSSION: Stripe 82 Science Verification plans [Fuyan, Elisabete, Andrew, Kyler]

- we need a target list that spans the full RA of Stripe 82, and with 6 degrees in Dec (i.e. extend 4 degrees in Dec to fit Taipan field)
- SDSS photometry and spectra + cross-match with Y1 input catalogue to check the 2MASS properties of the sources
- select targets with a range of air masses (assume August), RAs to get closer to the plane of the galaxy to check for star/galaxy separation
- 120 objects per field, 6 nights of SV, 24 fields/night = 144 field ~ 15,000 targets, some of which will be repeated
- we can select randomly from the catalogue (e.g. every nth object)
- assume r-limit of 17.5 (or 18?)
- [code to compute r-band fibre magnitude from the fibre spectra -> ask Matt Owers for his code]
- [flux calibration code from Millie]
- test increasing integration times; plot S/N as a function of magnitude and time (to test spectrograph performance); 15min, 30min, 1 hour on same targets, go as deep as r=18.
- take calibration frames for every set of 15min exposures; then compare the results with using calibration frames from the beginning of the night to test if we can get away with taking calibration frames only once per night
- if possible do these tests for three Taipan fields for more stats + the edges for airmass tests
- Baldry+2010 has colour cuts for GAMA for getting F stars for spectrophotometric calibration; these can be used as standards for testing the tiling code in this region; we should also get a catalogue of guide stars
- [astrometry: positional offsets between SDSS and 2MASS ~0.3 arcsec, needs to be corrected for pointing of starbugs]

- all these tests can be done in 2 or 3 nights
- include tests of ancillary targets: QSOs, low surface brightness galaxies, radio sources with no optical counterparts (from FIRST)
- write up detailed plan by Feb 1st

DISCUSSION: SPT Deep Field Science Verification plans [Ned and others]

- this field will be used to test survey operations more than performance
- full survey depth over 4 pointings
- new science! redshifts of radio sources, BPT diagnostics, spectroscopic SFRs
- write up detailed plan by Feb 1st

DISCUSSION: Taipan Y1

- cosmology (H0): 400,000 targets, 98% completeness, 4200 tiles, can be done in 120 nights [note: this needs to be double-checked properly with tiling/scheduling -> need survey simulator for that!]
- peculiar velocity sample: ~100,000 targets, with probably repeat observations: is this contained in the H0 sample or do we need to add an extra selection [Ned & Christina to work on it]
- galaxy evolution: pilot Wallaby fields will be very small and quick to to; we can add the KiDS survey. Strengths: good morphology and multi-wavelength ancillary data (VISTA, GALEX, WISE); 1500 sq. deg. can be done in 800 tiles according to Ned [TBC]. We need to figure out what competitive science can be done with this dataset!
- guide stars: no need to worry about it for now according to Kyler; FunnelWeb are planning to use APASS
- standard stars: we still don't have a catalogue for those. Plan A: SkyMapper shallow which we should have access to early next year if all goes well; plan B: PanSTARRS (it's missing the southern part below -30 deg, but we could start where we have it while waiting for SkyMapper)

Miscellaneous:

- Kyler showed us a nice tool to manage tasks as a team called "trello" [Elisabete: look into setting up one of those for the Taipan team]
- Taipan white paper: work on it next year, model after the SAMI paper (Croom+2012)
- **Next meeting:** Feb 1st — 3rd (at AAO); option for people to stay the full week and work
- Taipan team policy will be ready by Feb. meeting [Elisabete, Andrew]
- define list of tasks where team member contribution is needed (below):

TASKS (ordered by priority):

Note: names in parenthesis are the leads of these efforts/discussions, but we need more team members to work on each task!

1. **[by Feb 1st]** Survey simulator (including routing and allocation of fibres); use 10x10deg sandbox area [Ned, Elisabete, Chris B.]

2. **[by Feb 1st]** Finalize Science Verification plans (written up by Feb.1st) [SPT: Ned; Stripe 82: Fuyan]

3. **[by Feb 1st]** Final versions of the input catalogues for Y1 (write-up Taipan Y1 survey plan by Feb. 1st) [H0: Chris B.; PV: Christina; galaxy evolution: Ned]

4. **[by Feb 1st]** Database structure (including FITS header requirements)
   4.1. define requirements [Exec]
   4.2. look into OzDES structure, talk to Chris Lidman [Andrew]
   4.3. implementation [Sam, Andrew]

5. Archive schema, naming conventions etc [Andrew, Ned, Matthew, Sam]

6. 2dFDR/Taipan [talk to Mike Birchall at AAO: Andrew]

7. Automated quality control check (2dFDR? MARZ?) [Matthew, Sam, Mike]

8. Scheduler [contact AAO: Andrew]

9. Coordinate with FunnelWeb to see what type of effort is available from them [Andrew, Elisabete, Kyler]

10. Follow instrument checks with instrumentation team [Kyler, Exec]

11. Allocation of sky fibres [Nuria]

12. Draft white paper [Elisabete, Andrew]