Optimising a spectroscopic training sample with 4MOST

Isobel Hook and Jon Carrick

4MOST-TiDES team
TiDES: Time Domain Extragalactic Survey

- Everywhere 4MOST points, there will be known transients
- TiDES will use ~2% of 4MOST low-res fibres in extragalactic fields
- Three goals
  - Redshifts of SN host galaxies, after the SN has faded – up to 50,000
  - Spectra of live transients: classification, rare transients - up to 30,000
  - AGN reverberation mapping
- Allocation of 250,000 fibre hours - survey design underway now
- See Swann et al 2019

TiDES founding institutes

[Logos of the founding institutions]
This work

• Main science goal: Type Ia SN cosmology

• How should we use 4MOST to create the best possible training sample for photometric classification?

• For now, we are considering the final classification, not early classification for triggering follow-up
Capabilities of 4MOST

• Fixed exposure time:
  - TiDES has 250,000 fibre-hours
  - Baseline 120 min per field visit
  - Longer in 4MOST deep fields

• A 4MOST-only training sample will be magnitude limited to $r(AB) \sim 22.5$ mag

Typical spectra in one hour of exposure, Swann et al 2019
TiDES spectral success rate

Completeness = Success rate \textit{for those transients attempted}

Success defined as $S/N > 10$ per 15A bin, averaged over 4500 - 8000 A

From the 20/5/19 4MOST Exposure Time Calculator run
Classification method

- *snmachine* (Lochner et al. 2016)
- Using wavelets feature extraction
- Supernova Photometric Classification Challenge (SPCC) data.
- 21,000 objects with DES-like lightcurves, 1103 training objects with known classes
- Future: PLAsTiCC with new version of *snmachine* (late 2019?) + hopefully other classifiers to compare performance

SPCC Type Ia SN light curve fit with wavelets
Method

• Different training samples selected based on...
  • Magnitude limit, class balance, extra faint objects (from another telescope or data augmentation)

• Metric: ROC curves and area under Curve (AUC)
  • general performance of classifier for binary classification of Ia vs CC

• We used “spectroscopic z” throughout
  • Assumes that for Ia Hubble diagram, objects have at least a host spectrum
Representative training sample

• 500 training objects selected randomly (class and magnitude representative)
Representative training sample

- 500 training objects selected randomly (class and magnitude representative)

Mean AUC for 1103 training objects = 0.958
Mean AUC for 500 objects = 0.943
Magnitude-limited training sample

- 500 training objects sampled randomly following spectral completeness curve with 50% max
Magnitude-limited training sample

- 500 training objects sampled randomly following spectral completeness curve with 50% max

AUC average = 0.709
Balancing classes

- A magnitude-limited sample is not balanced in terms of classes
- Can create a magnitude limited, class-balanced set
  - Proportions of three classes forced to match the test set
  - Kept total number of training objects the same
Magnitude-limited training sample \( (r < 22.5) \)

Classes not balanced

Classes balanced

Conclusion – balancing classes doesn’t help much. Magnitude limit affects the performance.
Adding fainter objects

- Adding one extra supernovae per type (Ia, Ibc, II) per 0.5 magnitude bin – 30 total
Adding fainter objects – results before

AUC average = 0.709
Adding fainter objects – results after

AUC average = 0.727
Adding fainter objects

• Adding a few faint objects helps a bit
  • Depends on algorithm and on the exact set of objects chosen.

• But also implies longer exposure times (not easy with 4MOST) or using different spectroscopic facilities

• Try data augmentation (preliminary)
  • Using Avocado method (Boone 2019)
  • Resamples training lightcurves at new redshift (redshifted bandpasses), dims by DM and adds noise(*)
Data augmentation (preliminary)

- Training data augmented x50, and extends to higher redshift
Adding fainter objects – results after

AUC average = 0.727
Data augmentation – results after

AUC average = 0.872
## Results summary

<table>
<thead>
<tr>
<th>Training sample type</th>
<th>N training</th>
<th>4MOST exp time / $10^6$ s</th>
<th>Mean AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representative</td>
<td>500</td>
<td>343.4</td>
<td>0.943</td>
</tr>
<tr>
<td>Mag limited</td>
<td>498</td>
<td>7.9</td>
<td>0.709</td>
</tr>
<tr>
<td>Faint added</td>
<td>526</td>
<td>147.3</td>
<td>0.727</td>
</tr>
<tr>
<td>Augmented</td>
<td>498 (aug. to 24636)</td>
<td>7.9</td>
<td>0.872</td>
</tr>
</tbody>
</table>

Preliminary conclusion:
Data augmentation very promising but doesn’t capture everything (Also potential biases)
Conclusions

• 4MOST-TiDES will accumulate the largest spectroscopically-confirmed Ia sample for cosmology to date
• The live transient sample will also act as a training sample for photometric classification of other LSST transients, but with a magnitude limit that affects performance
• To address this, we must either reallocate exposure time, or make use of other spectroscopic facilities, or consider data augmentation methods
• ...or accept a reduced sample size
• Paper in preparation (Carrick et al)