A cooperative approach among SED and ML methods

Stefano Cavuoti\textsuperscript{1}, Massimo Brescia\textsuperscript{1}, Valeria Amaro\textsuperscript{2}, Civita Vellucci\textsuperscript{2}, Giuseppe Longo\textsuperscript{2}

\textsuperscript{1} INAF - Observatory of Capodimonte, Naples
\textsuperscript{2} University of Naples Federico II, Naples

Cavuoti et al. 2016, Submitted to MNRAS
Photo-z main open issues recap

1. General accuracy of photo-z’s in the Knowledge Base ranges;
2. Reliable PDFs for photo-z’s predicted by empirical methods;
3. **Virtuous cooperation among theoretical and empirical methods**;
4. Extension of accurate photo-z estimation beyond the spectroscopic range imposed by KB for empirical methods;
5. Combination of PDFs obtained by ML and SED fitting methods.
In this work we perform a comparison between five different photo-z techniques applied to the same KiDS dataset:

1. **MLPQNA** (Multi Layer Perceptron with Quasi Newton Algorithm);
2. **RF** (Random Forest);
3. **LEMON** (LEvenberg-Marquardt Optimization Network);
4. **Le Phare** SED template fitting;
5. **BPZ** (Bayesian Photometric Redshift model).

Then we propose a combination of different methods that provides an improvement in the accuracy of the final estimates.
The Data (KiDS DR2) pre-processing

We used the KiDS DR2 (de Jong et al. 2015), photometry with SDSS and GAMA spectra as KB

- excluded objects with low photometric quality (i.e. with flux error higher than one magnitude);
- removed all objects having at least one missing band (or labeled as Not-a-Number or NaN);
- selected objects with IMA FLAGS equal to zero in the g, r and i bands (i.e. sources that have been flagged because located in proximity of saturated pixels, star haloes, image border or reflections, or within noisy areas). The u band is not considered since the masked regions relative to this waveband are less extended than in the other three KiDS bands.

The final KB consisted of 15,180 training and 10,067 test objects
In a recent paper (Cavuoti et al. 2015) we provided a catalogue of photometric redshifts for about 1 million of KiDS galaxies, using MLPQNA.
Robustness Experiments

- $EX_{\text{clean}}$: fully corrected photometry
- $EX_{\text{ext}}$: corrected by extinction but with a residual offset
- $EX_{\text{off}}$: without the offset but not corrected by extinction
- $EX_{\text{no}}$: not corrected by extinction and with a residual offset

Highlights

- SED fitting methods less accurate than ML models;
- Residual offset has a non-negligible impact on ML methods also;
- ML methods robust to reddening;
- Le Phare more robust than BPZ to reddening;
- The lower impact of offset and reddening on estimators $\sigma_{68}$ and NMAD is justified by their lower dependence from outliers;
- More in general, the most relevant affecting factors are residual offset and outliers

<table>
<thead>
<tr>
<th></th>
<th>EXP</th>
<th>MLPQNA</th>
<th>LEMON</th>
<th>RF</th>
<th>Le Phare</th>
<th>BPZ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bias</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0010</td>
<td>0.0121</td>
<td>0.0289</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{ext}}$</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0012</td>
<td>0.0183</td>
<td>0.0393</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{off}}$</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0010</td>
<td>0.0158</td>
<td>0.0405</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{no}}$</td>
<td>0.0009</td>
<td>0.0010</td>
<td>0.0012</td>
<td>0.0225</td>
<td>0.0496</td>
<td></td>
</tr>
<tr>
<td><strong>$\sigma$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.026</td>
<td>0.026</td>
<td>0.029</td>
<td>0.065</td>
<td>0.127</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{ext}}$</td>
<td>0.028</td>
<td>0.028</td>
<td>0.030</td>
<td>0.079</td>
<td>0.218</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{off}}$</td>
<td>0.026</td>
<td>0.026</td>
<td>0.029</td>
<td>0.066</td>
<td>0.142</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{no}}$</td>
<td>0.028</td>
<td>0.028</td>
<td>0.030</td>
<td>0.079</td>
<td>0.222</td>
<td></td>
</tr>
<tr>
<td><strong>$\sigma_{68}$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.018</td>
<td>0.018</td>
<td>0.021</td>
<td>0.041</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{ext}}$</td>
<td>0.021</td>
<td>0.020</td>
<td>0.023</td>
<td>0.048</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{off}}$</td>
<td>0.018</td>
<td>0.019</td>
<td>0.021</td>
<td>0.041</td>
<td>0.045</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{no}}$</td>
<td>0.021</td>
<td>0.020</td>
<td>0.023</td>
<td>0.049</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td><strong>NMAD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.018</td>
<td>0.018</td>
<td>0.021</td>
<td>0.038</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{ext}}$</td>
<td>0.020</td>
<td>0.020</td>
<td>0.022</td>
<td>0.044</td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{off}}$</td>
<td>0.018</td>
<td>0.018</td>
<td>0.021</td>
<td>0.037</td>
<td>0.033</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{no}}$</td>
<td>0.020</td>
<td>0.020</td>
<td>0.022</td>
<td>0.044</td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td><strong>% Outliers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.31</td>
<td>0.30</td>
<td>0.40</td>
<td>0.89</td>
<td>2.18</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{ext}}$</td>
<td>0.34</td>
<td>0.35</td>
<td>0.42</td>
<td>2.51</td>
<td>3.83</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{off}}$</td>
<td>0.31</td>
<td>0.29</td>
<td>0.39</td>
<td>1.12</td>
<td>3.21</td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{no}}$</td>
<td>0.33</td>
<td>0.36</td>
<td>0.36</td>
<td>2.63</td>
<td>4.37</td>
<td></td>
</tr>
</tbody>
</table>
First Results \( (EX_{\text{clean}} \text{ experiment type}) \)
Spectral-type classification based on Le Phare without bounding the fitting with any kind of redshift

The evident performance variation for different morphological types, induced us to explore the possibility to combine the methods, by exploiting Le Phare spectral-type classification to specialize ML methods to predict photo-z's for objects belonging to a single spectral class.

**Recap of \( \text{EX}_{\text{clean}} \) experiment type**
1. Derive traditional photo-z’s with all methods;
2. Use Le Phare bounded with spec-z’s to obtain a reference classification;
3. Use Le Phare bounded with photo-z’s to perform a series of classifications;
4. Identify the best classification using as ground truth the reference classification (step 2);
5. Perform a photo-z regression by training MLPQNA on separated subsets specific for each class;
6. Recombine the output.
Step 1 - Usual Redshifts

<table>
<thead>
<tr>
<th>EXP</th>
<th>MLPQNA</th>
<th>LEMON</th>
<th>RF</th>
<th>Le Phare</th>
<th>BPZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0010</td>
<td>0.0121</td>
<td>0.0289</td>
</tr>
<tr>
<td>$\sigma$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.026</td>
<td>0.026</td>
<td>0.029</td>
<td>0.065</td>
<td>0.127</td>
</tr>
<tr>
<td>$\sigma_{68}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.018</td>
<td>0.018</td>
<td>0.021</td>
<td>0.041</td>
<td>0.039</td>
</tr>
<tr>
<td>NMAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.018</td>
<td>0.018</td>
<td>0.021</td>
<td>0.038</td>
<td>0.031</td>
</tr>
<tr>
<td>% Outliers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$EX_{\text{clean}}$</td>
<td>0.31</td>
<td>0.30</td>
<td>0.40</td>
<td>0.89</td>
<td>2.18</td>
</tr>
</tbody>
</table>
Step 2 and 3 - Le Phare Classification (bounded by photo-z’s)

Confusion Matrices

Scd and SB classes are always well classified

E/S0 and Sab are classified better by ML redshifts

E type is classified better by Le Phare and RF redshifts

Therefore, RF redshifts are the best candidate (for such purpose)
Step 4 - Find the best Classification

Reference classification

Le Phare photo-z
MLPQNA photo-z
LEMON photo-z
RF photo-z
BPZ photo-z

Spec-z
Le Phare classifier

Best classification

Normalized Confusion Matrix RF

True label

Predicted label

E
E/S0
Sab
Scd
SB

E
E/S0
Sab
Scd
SB
Step 5 and 6 - Improved Redshifts by recombination
Conclusions

The proposed workflow, involving different methodologies by mixing for the first time in a single collaborative framework SED fitting and machine learning models, is able to improve the photo-z prediction accuracy by ~10%.

The performance are strongly depending on the class definition; therefore on the SED models selected and on the SED fitting setup.

When a proper classification is provided, the photo-z’s produced by ML methods would benefit.
Photo-z main open issues recap

1. General accuracy of photo-z’s in the Knowledge Base ranges;
2. Reliable PDFs for photo-z’s predicted by empirical methods;
3. **Virtuous cooperation among theoretical and empirical methods**;
4. Extension of accurate photo-z estimation beyond the spectroscopic range imposed by KB for empirical methods;
5. Combination of PDFs obtained by ML and SED fitting methods.

Status

1. Empirical methods outperform theoretical ones.
2. METAPHOR seems to be a good candidate to solve the problem.
3. **The proposed combination method is a good starting point.**
4. **STILL OPEN;**
5. **STILL OPEN.**