Challenges in Panchromatic Modelling with Next Generation Facilities
Proceedings IAU Symposium No. 341, 2018
M. Boquien, E. Lusso, C. Gruppioni & P. Tissera, eds.

doi:10.1017/S1743921319002515

Predicting the global far-infrared emission of galaxies

Wouter Dobbels and Maarten Baes

Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281, B-9000 Gent, Belgium
email: wouter.dobbels@ugent.be

Abstract. Dust absorbs stellar emission and reradiates this energy in the far-infrared (FIR). FIR observations hence give us a direct view of the dust, and allow us to study its properties. Unfortunately, FIR observations are only available for a small subset of galaxies. In this work, we estimate the global FIR emission from global UV-NIR observations. We show that a machine learning method clearly outperforms a SED modelling approach. For each galaxy, we not only predict the FIR flux across the 6 Herschel bands, but also estimate individual uncertainties. We inspect the worst predictions, and investigate how the machine learning predictor generalizes on new data. Our predictor can be used as a virtual observatory, which is especially useful now that there is still no confirmed next-generation FIR telescope.

Keywords. infrared: galaxies, infrared: ISM, galaxies: ISM

1. Introduction

When studying galaxies through broadband photometry, two main components arise. At ultraviolet (UV) to near-infrared (NIR) wavelengths, the stellar emission dominates. This emission is partially absorbed by dust, heating the dust. The dust then emits thermally at longer wavelengths. Very small grains are stochastically heated, and dominate at mid-infrared (MIR) wavelengths. Larger grains are in thermal equilibrium, and they emit at FIR wavelengths. We can learn some things about the dust through UV-NIR attenuation (e.g., Charlot & Fall 2000; De Geyter *et al.* 2014; Leja *et al.* 2017), however there is some degeneracy between attenuation, stellar age and metallicity (Bell & de Jong 2001). The FIR can be used to break this degeneracy, and gives a direct, more accurate estimate of properties such as dust mass and effective temperature (see Conroy 2013). By assuming energy balance, the total IR luminosity L_{IR} equals the total attenuated at shorter wavelengths. Hence, FIR observations are not only useful to study dust, but help constrain other properties as well (Leja *et al.* 2017).

Unfortunately, the FIR is hard to observe. Ground observations are not possible, cooling is required to limit the thermal noise of the instrument itself, and the angular resolution is a lot worse than for optical telescopes. The Herschel Space Observatory stopped operations in 2013 after cooling ran out, and a follow-up mission is still not confirmed. Meanwhile, excellent optical and NIR telescopes are upcoming, including JWST, LSST, Euclid and WFIRST. For many galaxies, both nearby and at larger redshift, we have UV-NIR observations but no FIR observations.

The goal of this work is to predict the FIR from UV-NIR observations. We know the two are related through the dust energy balance. However, even if the total attenuation can be recovered perfectly, it only constrains L_{IR} and not the shape of the infrared SED. Dust emission and attenuation are usually modelled independently, but this of course doesn't mean that the two are unrelated.

To achieve this goal, our main result uses machine learning techniques. This ensures that we learn an optimal mapping from input (UV-NIR) to output (FIR), given a large enough training dataset. However, we start with a more traditional SED modelling approach.

2. Dataset and SED modelling

In order to test our methods, we require a set of galaxies with UV to FIR observations. For this, we use two samples. The first is H-ATLAS DR1 (Eales *et al.* 2010; Valiante *et al.* 2016), which covers 3 equatorial fields also observed by the GAMA survey (Driver *et al.* 2011). After limiting these to $z < 0.1$, this sample contains 3 618 galaxies. The second sample is DustPedia (Davies *et al.* 2017; Clark *et al.* 2018), a sample of local ($z < 0.01$) galaxies observed by Herschel. After removing galaxies with contamination and galaxies with insufficient datapoints (minimum of 4 UV-NIR and 3 FIR observations), we have 715 galaxies remaining. For now, the two datasets are combined into a total sample of 4 333 galaxies.

The FIR can be estimated from UV-NIR through Bayesian SED modelling, using stellar population synthesis (SPS). For this, we used the SED fitting code CIGALE (Boquien *et al.* 2018), with a grid of nearly 80 million models (Nersesian et al., in prep). The models are only fit to UV-NIR observations. From this, we extract a Bayesian estimate of each of the 6 Herschel bands (70 μm – 500 μm).

We could compare these predictions against the raw Herschel observations. However, the raw data do not always lead to a realistic SED, and include missing fluxes, negative fluxes (due to sky subtraction) and some outliers (resulting in unrealistic peaks or troughs). We can do better by making use of all information: a full UV-FIR SED fit. For this we use a similar setup as the short wavelength CIGALE fit, but this time using UV-FIR data. We extract the Bayesian estimates for the 6 Herschel bands, which are now much better constrained due to the corresponding observed data points. These serve as our ground truth FIR fluxes: a Bayesian estimate for a realistic SED which makes use of all available datapoints.

We can now compare the predicted values (from the UV-NIR CIGALE fit) to the ground truth values (from the UV-FIR CIGALE fit). Instead of directly comparing the fluxes, we first divide by the WISE 3.4 μm flux, in order to remove the dependency on distance and total luminosity. Then, we take the base 10 logarithm due to the large dynamic range. The comparison of these log normalized fluxes ($\log(F_{\rm FIR}/F_{3.4})$) is shown in Fig. 1 for each of the 6 Herschel bands.

We find a total root mean squared error (RMSE) of 0.39 dex across the 6 bands. There is a clear correlation between prediction and target, but there are definitely some biases, especially at longer wavelengths. There are two main problems. First: it is hard to estimate the total absorbed (and hence emitted) energy budget from UV-NIR wavelengths, due to the degeneracy discussed earlier. Second: even when this energy budget is known, we don't know the shape of the FIR, since the emission parameters are varied independently of the absorption parameters. The result of these predictions depend on the grid of models chosen. Leja *et al.* (2017) show that better predictions are possible by using more constrained priors. However, this choice of prior depends on the sample, and limits the flexibility.

3. Machine learning

Due to the model limitations described above, it is possible to do better. Machine learning ensures that we learn an optimal mapping between the input (UV-NIR fluxes) and output (FIR fluxes), at least for the training set. The method can then be evaluated

https://doi.org/10.1017/S1743921319002515 Published online by Cambridge University Press

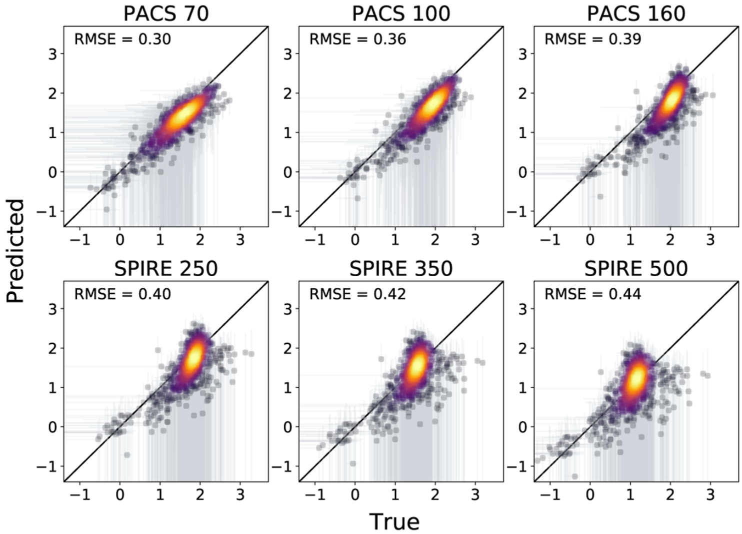

Figure 1. Comparing the SED modelling predictions against the ground truth. Both predicted and true values are Bayesian estimates of the FIR fluxes from CIGALE, and show the log normalized fluxes ($\log(F_{\mathrm{FIR}}/F_{3.4})$). However, the predicted values only make use of UV-NIR data, while the true values make use of UV-FIR data.

on unseen data: the test set. We (randomly) set aside 25% of the data as a test set. For the input, we again have problems when using the raw fluxes. In particular, machine learning has trouble using missing values and large uncertainties in the input. Hence, we use the UV-NIR CIGALE fit and extract the Bayesian estimates of the 14 UV-NIR fluxes that were observed. Of course, no Herschel observations are used in order to make a prediction. We use a 3-layer neural network with an Adam optimizer (Kingma & Ba 2014), using warm restarts. We pick the model that performed best on a cross-validation set (split from the training set).

The predictions on the test set are shown in Fig. 2. The total RMSE is now 0.20 dex, compared to 0.39 dex for the SED modelling. The result does not show significant biases. Note that we also estimated uncertainties on the predictions. For this we used the methodology of Gurevich & Stuke (2017): we train a second neural network as an uncertainty estimator, by optimizing a joint loss function. This allows us to identify which predictions can be trusted and which can not. When inspecting the worst predictions, we see that these can often be explained. Sometimes, the WISE 12 μm and WISE 22 μm bands are missing, and the Bayesian estimate of these values can mislead the neural network. Some galaxies also have unusual SED shapes, with possible AGN contribution, and did not resemble the training set.

Since machine learning tends to only work on samples that resemble the training set, we investigated how the algorithm performs on different data. When testing on a H-ATLAS sample of $0.1 < z < 0.5$ (training was for $z < 0.1$), we find a total RMSE of 0.27. This is a satisfying result, meaning that there is some room for extrapolation. When training only on H-ATLAS and testing on DustPedia, the total RMSE is 0.40. There are several explanations for this rather large RMSE value. First of all, the two data sets were each reduced in their own way. Moreover, DustPedia is a more local sample and contains less strict criteria on FIR S/N. Especially at longer wavelengths, we find that DustPedia includes galaxies with lower FIR fluxes than available in the H-ATLAS set. Keeping these differences in mind, the neural network also produces satisfying results. The individual

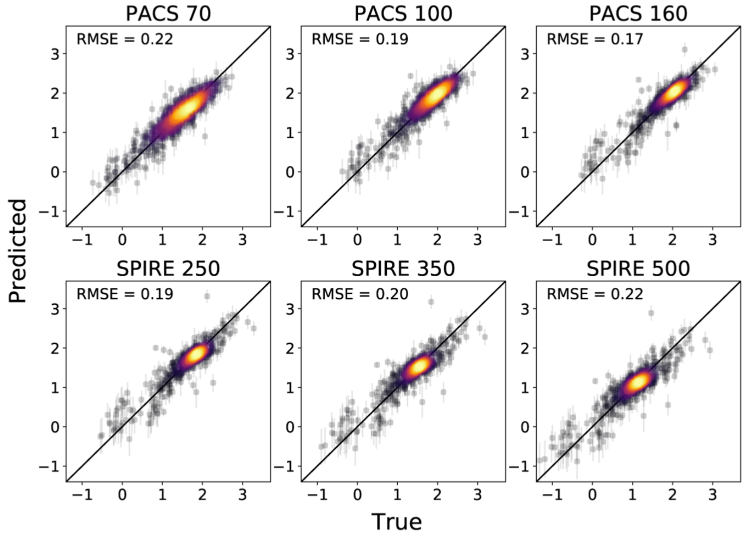

Figure 2. Similar to Fig. 1, but this time showing the neural network predictions against the Bayesian UV-FIR ground truth.

outliers can be inspected, and again it is possible to see in which scenarios the prediction is not trustworthy.

Overall, this neural network gives us an interesting tool to study the FIR. It is possible to explore which UV-NIR properties lead to a certain FIR shape, without being limited to a sample of real galaxies: we can predict for arbitrary UV-NIR input. Our model clearly outperforms SED modelling approaches. Through a predicted uncertainty, we know when the predictions can not be trusted (e.g. when there is large uncertainty on the input). Since we are able to predict a full FIR spectrum from UV-NIR, we are also able to predict related dust properties, such as dust mass and temperature. Since there are currently no confirmed next generation FIR telescopes, we hope these predictions can exploit the excellent upcoming optical/NIR telescopes in order to bridge this gap.

References

Bell, E. F. & de Jong, R. S. 2001, *ApJ*, 550, 212
Boquien, M., Burgarella, D., Roehlly, Y., Buat, V., Ciesla, L., Corre, D. Inoue, A. K., Salas, H., *et al.* 2018, arXiv:1811.03094
Charlot, S. & Fall, S. M. 2000, *ApJ*, 539, 718
Ciesla, L., Boselli, A., Elbaz, D., Boissier, S., Buat, V., Charmandaris, V. Schreiber, C., Béthermin, M., *et al.* 2016, *A&A*, 585, A43
Clark, C. J. R., Verstocken, S., Bianchi, S., Fritz, J., Viaene, S., Smith, M. W. L., Baes, M., Casasola, V., *et al.* 2018, *A&A*, 609, A37
Conroy, C. 2013, *ARAA*, 51, 393
Davies, J. I., Baes, M., Bianchi, S., Jones, A., Madden, S., Xilouris, M., Bocchio, M., Casasola, V., *et al.* 2017, *PASP*, 129, 044102
De Geyter, G., Baes, M., Camps, P., Fritz, J., De Looze, I., Hughes, T. M., Viaene, S., Gentile, G., *et al.* 2014, *MNRAS*, 441, 869
Driver, S. P., Hill, D. T., Kelvin, L. S., Robotham, A. S. G., Liske, J., Norberg, P., Baldry, I. K., Bamford, S. P., *et al.* 2011, *MNRAS*, 413, 971
Eales, S., Dunne, L., Clements, D., Cooray, A., De Zotti, G., Dye, S., Ivison, R., Jarvis, M., *et al.* 2010, *PASP*, 122, 499
Gurevich, P. & Stuke, H. 2017, arXiv:1707.07287
Jones, A. P., Köhler, M., Ysard, N., Bocchio, M., & Verstraete, L. 2017, *A&A*, 602, A46

Kingma, D. P. & Ba, J. 2014, arXiv:1412.6980
Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, *ApJ*, 837, 170
Valiante, E., Smith, M. W. L., Eales, S., Maddox, S. J., Ibar, E., Hopwood, R., Dunne, L., Cigan, P. J., *et al.* 2016, *MNRAS*, 462, 3146

Discussion

Calistro Rivera: How applicable is this technique at high redshifts, e.g. to predict ALMA fluxes? To which redshift does your sample go?

Dobbels: It's applicable as long as you have the training data for high-z and ALMA. In my work I'm limited to $z < 0.5$.

Goto: It seems like the predicted vs true slope is below 1?

Dobbels: By eye it looks like the predicted vs true lie on the 1:1 relation, but I can do a fit and hope that the slope is consistent with 1.

Buat: From the resulting SEDs you showed, it seems that the energy balance is not always preserved. Does this mean that your estimates are not completely physically consistent?

Dobbels: It is hard to estimate the total absorbed energy (more attenuation and older stars can be confused) from UV-NIR. The same happens when you fit CIGALE to UV-NIR: there will be significant scatter in the predicted L_{FIR} around the true values. Moreover, energy balance does not hold for a single viewing angle: edge-on and face-on have the same L_{FIR} (isotropic), but the edge-on is more attenuated.

Lee: What is the best strategy to choose the number of samples to train?

Dobbels: This is mainly driven by the test set, where we want decent statistics. I have 4000 total samples, and we used 1000 for testing (25%). I then look at the true vs predicted plot, and this seems to be well sampled. This can be quantified by taking a large number of bootstrap samples of the test set. If the bootstrap distribution of MSE is narrow, you have enough test samples. If you have 10 million total samples, 1% test size is probably enough.