APPENDIX A: PHYSICAL PARAMETER COMBINATION

In this section we verify if the independent derivation of the three physical properties, i.e. redshift, stellar mass, and SFR, bring to some unrealistic galaxies. Some examples of these objects are low-mass galaxies at high-z or low-mass quiescent galaxies, as both galaxy populations would be undetectable given the considered observational limits.

In Figure A1 we report the M_*−SFR plane of galaxies with S/N > 3 using the true values of the physical properties. In the same Figure we show, as example, the M_*−SFR plane using the physical properties derived with the DLNN with only *Euclid* filters for galaxies with S/N > 3. In addition, as a proxy for the star-formation main-sequence (MS, e.g. Noeske et al. 2007), we derive the median SFR value at different redshift, after selecting only star-forming galaxies, approximated as galaxies with $\log_{10}[\mathrm{SFR}/(\mathrm{M}_\odot\ \mathrm{yr}^{-1})] \geq -10.5$.

We can see that the derived SFR, stellar mass, and redshift are not completely unrelated, even if they have been derived independently. Indeed, the MS is present and its normalisation increases with redshift, as observed in the literature (e.g. Speagle et al. 2014; Schreiber et al. 2016) and as visible when considering the true values for the physical parameters. In addition, the number of quiescent galaxies ($\log_{10}[\mathrm{SFR}/(\mathrm{M}_\odot\ \mathrm{yr}^{-1})] < -10.5$) derived using DLNN or CNN varies between the 95-114% respect to their number in the input samples, indicating that the sSFR is recovered well enough for their identification.

Therefore, given the presence of the MS and the little variation on the recovered number of quiescent galaxies, we can conclude that there are no significant evidence of galaxies with nonphysical combinations of SFR, stellar mass, and redshift, even if these quantities are derived independently. We leave to future works the investigation of possible improvement when deriving the mentioned physical properties simultaneously with a single network as well as the full analysis of the derived MS and sSFR.

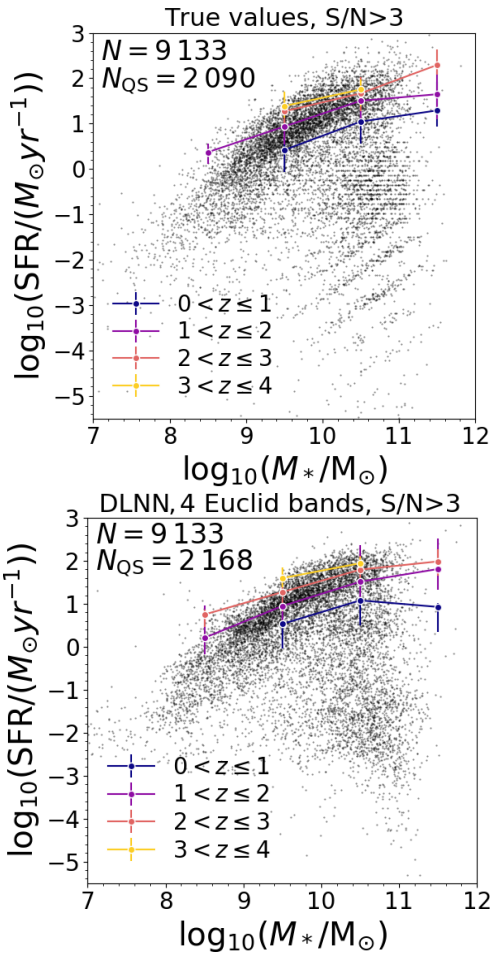

Figure A1. M_*−SFR plane for all galaxies with S/N > 3 considering the true values for redshift, stellar-mass, and SFR (*top*) and the values derived using a DLNN with four *Euclid* filters (*bottom*). Coloured lines indicate the median SFR of the star-forming galaxies (i.e. $\log_{10}[\mathrm{SFR}/(\mathrm{M}_\odot\ \mathrm{yr}^{-1})] \geq -10.5$) at $z = 1$ to 4. On the top left of each panel we report the total number of object in the sample and the number of quiescent galaxies (i.e. $\log_{10}[\mathrm{SFR}/(\mathrm{M}_\odot\ \mathrm{yr}^{-1})] > -10.5$).

APPENDIX B: SENSITIVITY ANALYSIS

To understand the importance of the different inputs considered in this work, we performed a sensitivity analysis (Guyon & Elisseeff 2003) of the parameter space used for redshift, stellar mass, and SFR estimations. For this test we considered the mock catalogue with S/N > 3 and nine input bands, i.e. the four *Euclid* filters and the u, g, r, i, and z ground-based ones.

In particular, we used the Light Gradient Boosting Model (LGBM; Ke et al. 2017) as the base regression model to perform different feature space analyses, using the knowledge base made by the training and testing sources and the related true targets. We optimised the model hyper-parameters through an automatic grid search for each of the regression use cases (redshift, stellar mass, and SFR), we also introduced six additional random features in the parameter space, made by a simple white noise. We then derived the informative contribution given by the input features by alternating four different methods, including:

- a feature importance calculation based on a standard tree-based method (i.e., XGBoost; Sagi & Rokach 2021);
- a Recursive Feature Elimination (RFE; Chen & Jeong 2007) that recursively fits a supervised algorithm considering a smaller sets of features. The excluded features are the ones that are considered less important according to the magnitude of some weights (e.g., the coefficients of linear models, or the feature importances for tree-based models).
- Boruta (Kursa et al. 2010), which is a wrapper-based technique for feature selection. In particular, we iteratively fitted a supervised algorithm (a tree-based model) on an extended version of the tabular data. The extended version, in each iteration, is composed of the original data with a horizontally attached shuffled copy of the columns. In each iteration, we maintained only the features that have a higher importance than the best of the shuffled features and are better than the expected random chance (using a binomial distribution).
- the SHAP method (SHapley Additive exPlanations; Lipovetsky & Conklin 2001; Štrumbelj & Kononenko 2013; Lundberg & Lee 2017), which was demonstrated to be effective on mitigating the effects in the selection of high-frequency or high-cardinality features. Taken from cooperative game theory, this method allows us to derive the contributions given by the presence, or absence, of the different input features.

We repeated the feature analysis for different training and testing by splitting seeds to mitigate randomness in data selection. We mixed the aforementioned methods through five different combinations, i.e. standard tree-based importance alone, RFE and standard tree-based

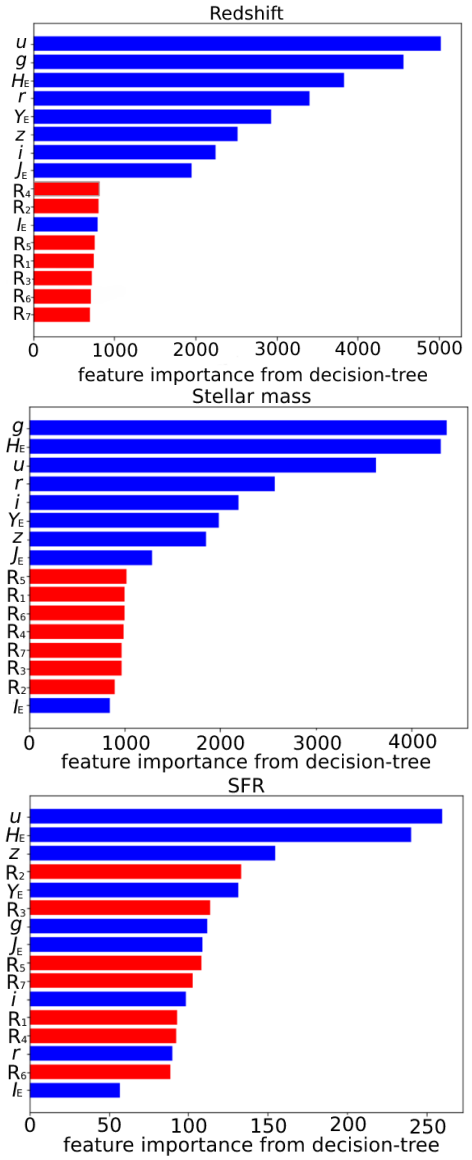

Figure B1. The feature importance calculation based on a standard tree-based method for redshift (*top*), stellar mass (*centre*), and SFR (*bottom*). Red bars indicate random features, while the blue ones are the nine considered filters.

importance, RFE and SHAP importance, Boruta and standard tree-based importance, and Boruta and SHAP importance.

In Figure B1 we show the results for the feature importance calculation based on the standard tree-based method as example, but the other bring to similar results. Overall, for all redshift, stellar mass, and SFR the optical bands have a high importance, while the I_E band is the less relevant. This may be caused by the wide wavelength range of the I_E filter that is already covered, but with higher spectral resolution, by the r, i, and z bands. It is however necessary to consider that the importance of the I_E filter may increase for galaxies, a minority in our mock catalogue, that are too faint to be observed in the single r, i, and z filters.

APPENDIX C: SED FITTING RESULTS

In this section we give more details on results derived with the SED fitting described in Section 4.

In particular, we report in Figure C1 the results for the redshift. Some of the most extreme outliers, when only the *Euclid* filters are as input, are galaxies between $z = 1$ and 2 that are wrongly classified as galaxies at $z = 4 - 5$. This is due to 4000-break that is wrongly identify by the the SED fitting code as the Lyman-break at 912, due to the limited wavelength coverage. However, when nine filters are as input, the improvement is not limited to these extreme outliers, but the general dispersion at $z < 2$ is reduced.

In Figure C2 we show the results for the stellar mass. As the stellar mass is derived in the same run as the redshift, a wrong redshift estimation will influence the stellar mass derivation. This explain the spread visible in the stellar mass derived when only *Euclid* filters are considered as input.

Finally, in Figure C3 we report the results for the SFR, as derived with SED fitting. The estimation largely improves when considering nine filters as inputs instead of four filters. This may be due to two reasons: i) a improvement in the redshift estimation, as mentioned before for the stellar mass; ii) the presence of the u-band filter which is more sensitive to SFR than the I_E band, being at shorter wavelengths.

REFERENCES

Chen X.-w., Jeong J. C., 2007, Proc. Sixth Int. Conf. Mach. Learn. Appl. (ICMLA 2007), Enhanced Recursive Feature Elimination. IEEE, Cincin- nati, OH, USA, p. 429
Guyon I., Elisseeff A., 2003, J. Mach. Learn. Res., 3, 1157
Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T.-Y., 2017, Adv. Neural Inf. Process. Syst., 30, 3149
Kursa M. B., Jankowski A., Rudnicki W. R., 2010, Fundam. Inform., 101, 271
Lipovetsky S., Conklin M., 2001, Appl. Stochastic Models Bus. Ind., 17, 319
Lundberg S. M., Lee S.-I., 2017, Adv. Neural Inf. Process. Syst., 30
Sagi O., Rokach L., 2021, Inform. Sci., 572, 522
Schreiber C., Elbaz D., Pannella M., Ciesla L., Wang T., Koekemoer A.,Rafelski M., Daddi E., 2016, A&A, 589, A35
Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15
Štrumbelj E., Kononenko I., 2013, Knowl. Inform. Syst., 41, 647

This paper has been typeset from a TEX/LATEX file prepared by the author.

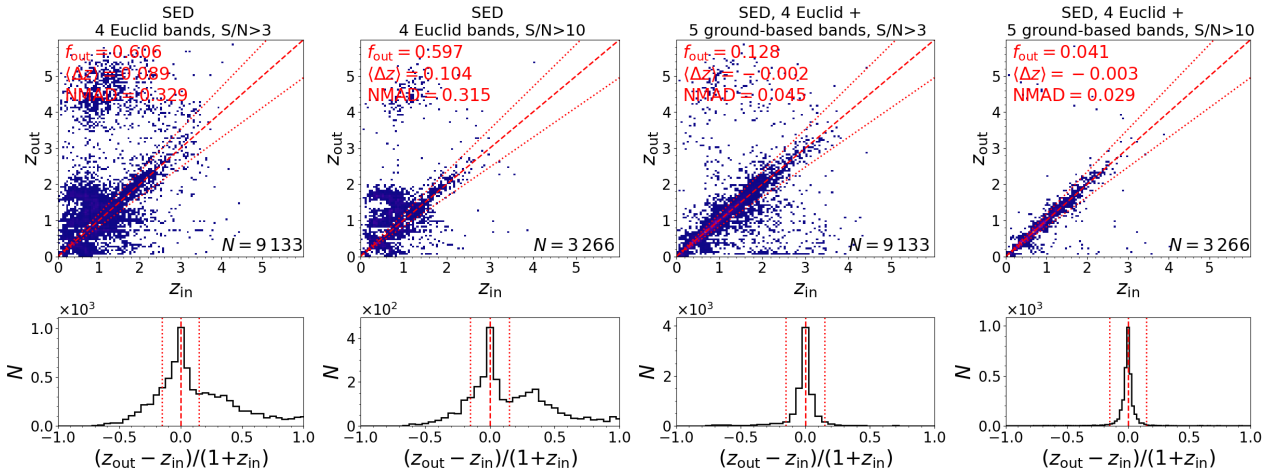

Figure C1. Same as Figure 4, but for the SED fitting.

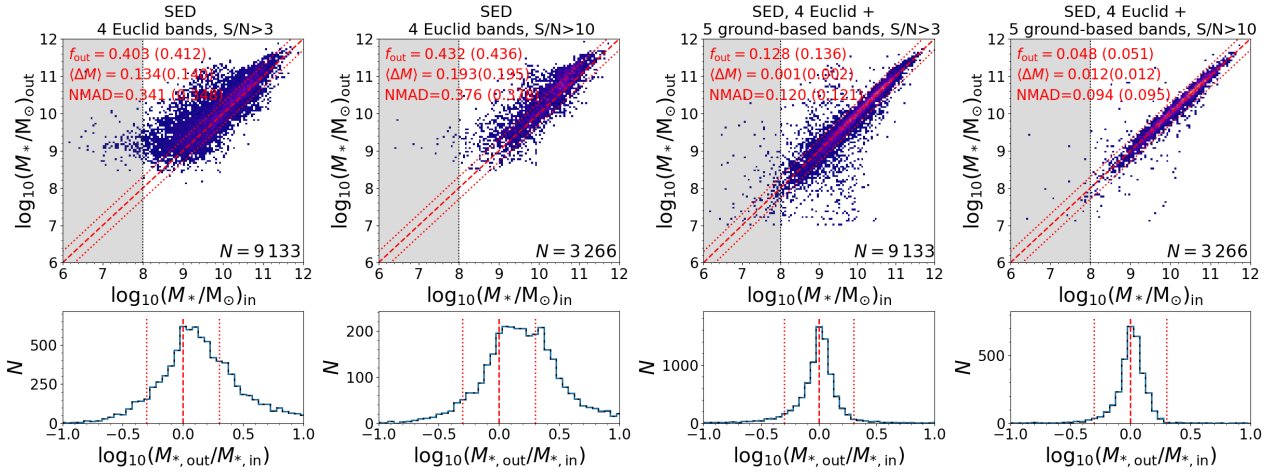

Figure C2. Same as Figure 7, but for the SED fitting.

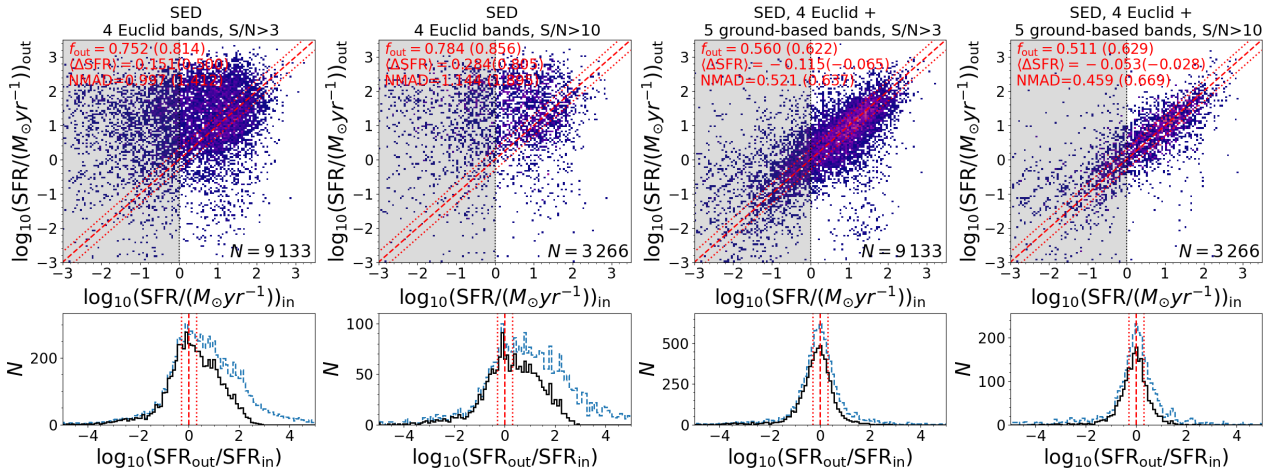

Figure C3. Same as Figure 14, but for the SED fitting.