USING FORS2 DATA TO IMPROVE TEMPLATES SELECTION FOR PHOTOMETRIC REDSHIFTS

JOSEPH CHEVALIER, SYLVIE DAGORET-CAMPAGNE

RÉZA ANSARI, JOHANN COHEN-TANUGI, MARC MONIEZ , ÉRIC NUSS

LSST-FRANCE, CC-IN2P3

DECEMBER 13, 2023

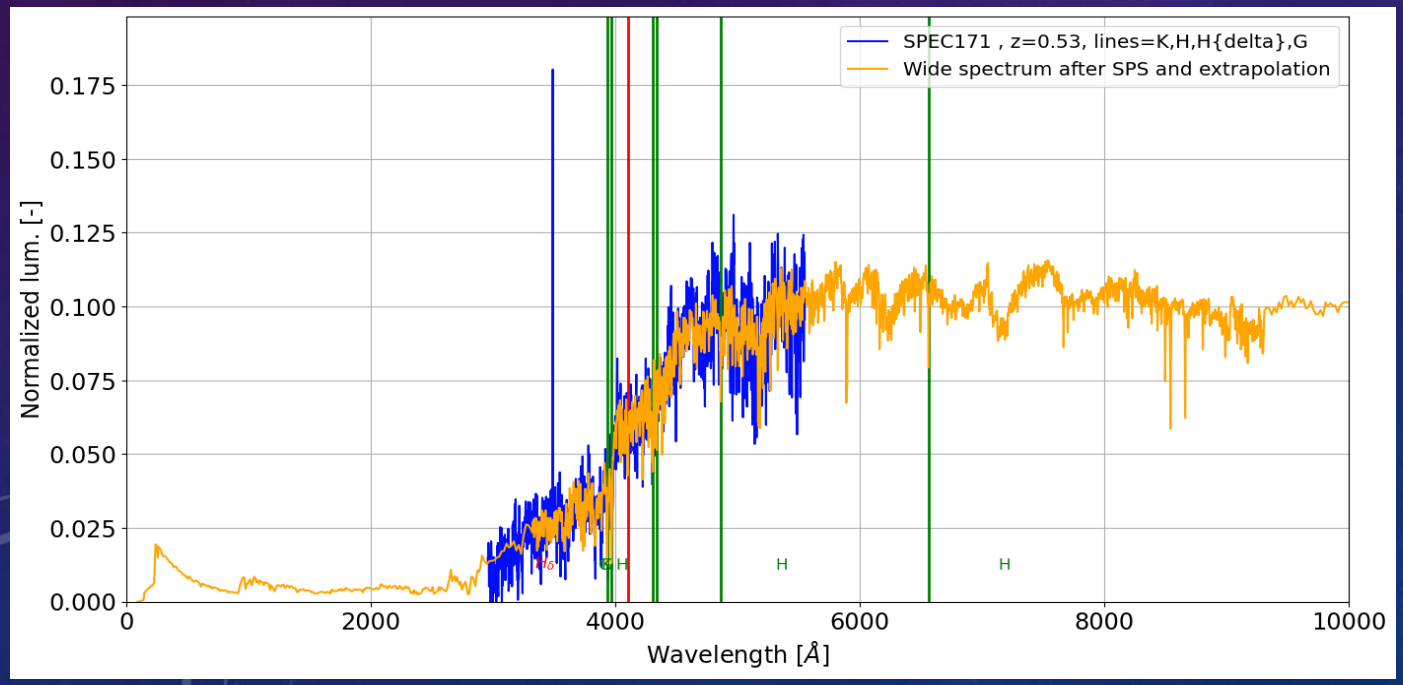
OUTLINE

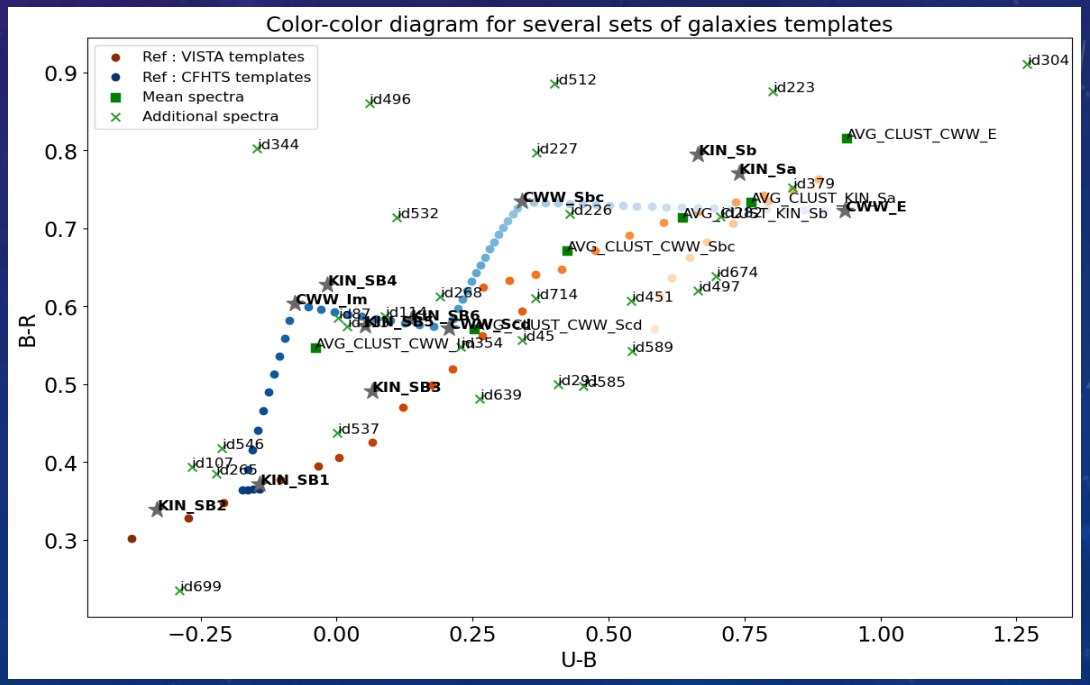
1. Get raw data : 550 narrow spectra from FORS2 at VLT
2. Extrapolate observations *(Stellar Population Synthesis)* → 550 wide spectra, from UV to IR
3. Select ≈ 30 as templates for photo-z estimation : *color indices, clustering, etc.*
4. Run photo-z (*LEPHARE++, in-house software*) **beyond point-estimates** :
    1. Identify best model at spectroscopic redshift
    2. Compare to the model and redshift selected IAW photometry
    3. Evaluate and understand the good *vs* bad templates
5. Is there a stellar population recipe for good templates?

OUTLINE

1. Get raw data : 550 narrow spectra from FORS2 at VLT
2. Extrapolate observations *(Stellar Population Synthesis)* → 550 wide spectra, from UV to IR
3. Select ≈ 30 as templates for photo-z estimation : *color indices, clustering, etc.*

LSST-France Grenoble June 2023





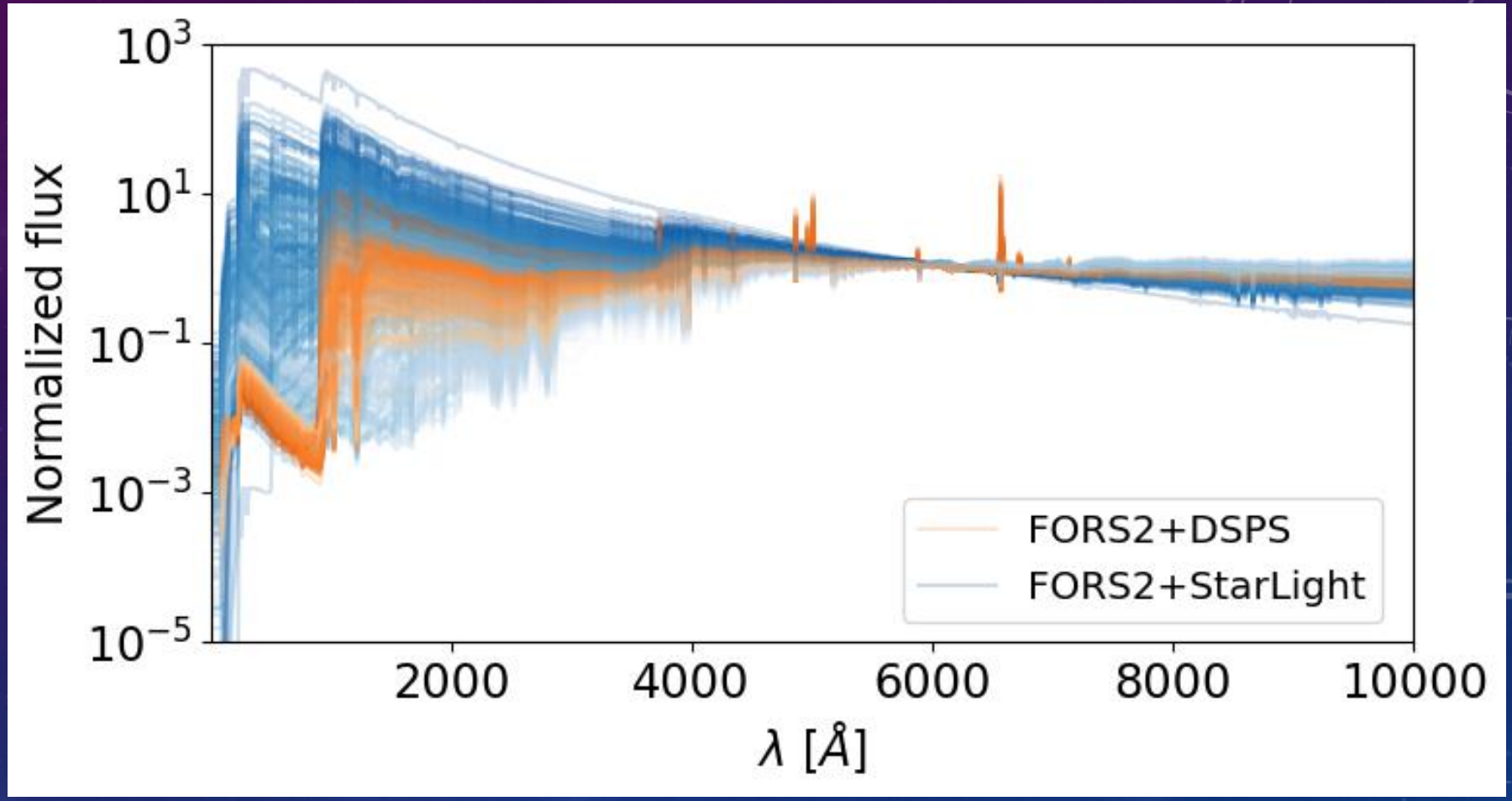
OUTLINE

4. Run photo-z (*LEPHARE++, in-house software*) **beyond point-estimates** :
    1. Identify best model at spectroscopic redshift
    2. Compare to the model and redshift selected IAW photometry
    3. Evaluate and understand the good *vs* bad templates

Focus of today's presentation

+ glimpse at potential results… if all things went perfectly !

OUTLINE

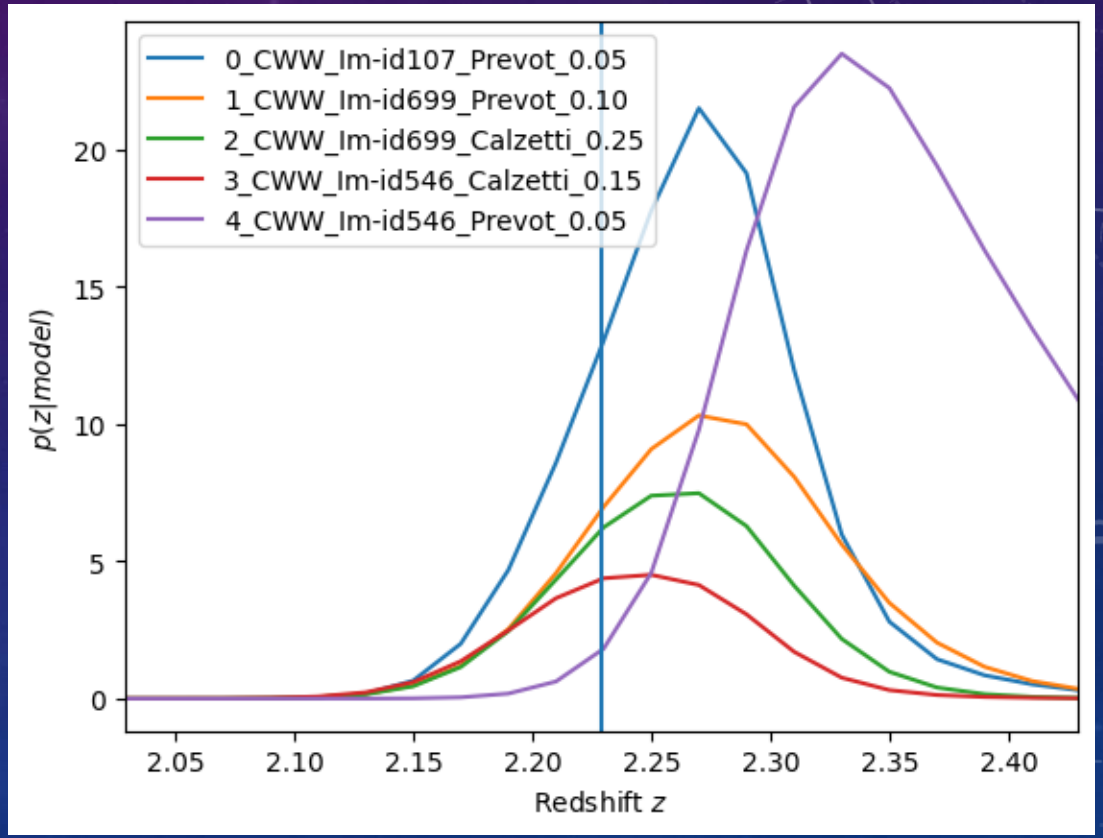


5. Is there a stellar population recipe for good templates?

BEST MODEL AT TRUE REDSHIFT

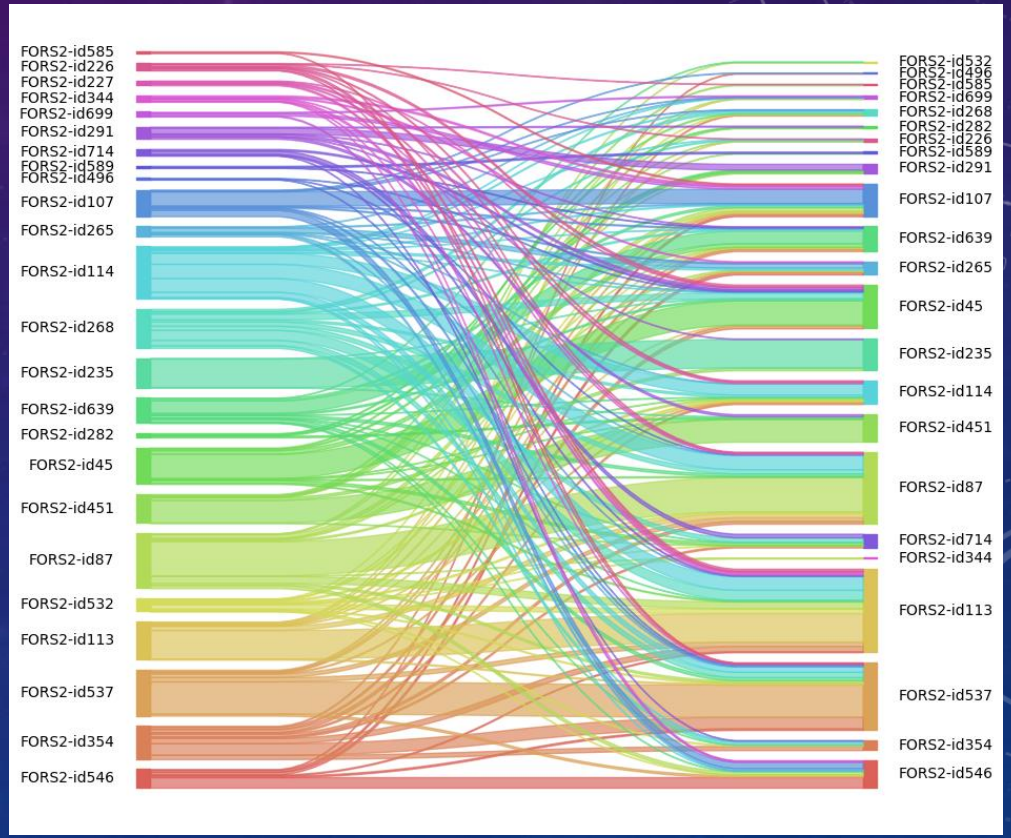
- Not a full photo-z estimation :
  - In LEPHARE++ : '`-ZFIX YES`' outputs the most-likely model (*i.e.* min(χ^2)) at specified redshift
  - In EmuLP : full posterior, evaluated at z_{spec} and highest probability wins !

- Would not necessarily output z_{spec} as the point-estimate for a given galaxy :
  - Mode, mean or median of p(z|model) ≠ z_{spec}
  - Degeneracies (dust *vs* redshift) → bimodal distributions



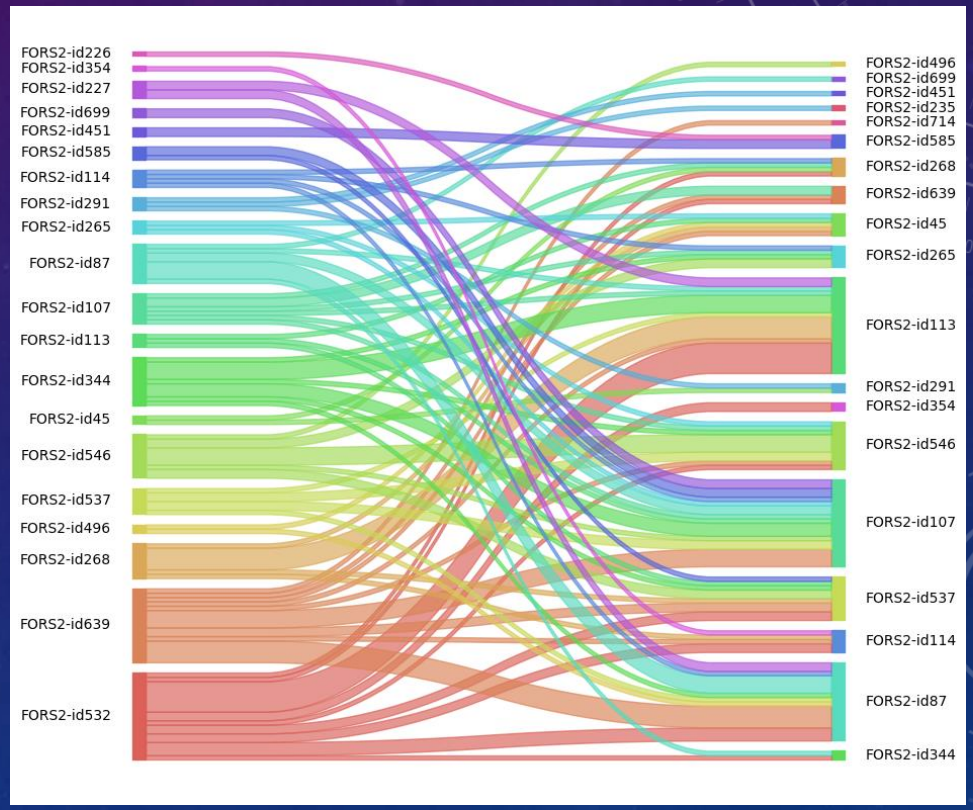
BEST MODEL AT TRUE REDSHIFT

- Comparison : best model @ z_{spec} =? model at min(χ^2)
  - 1000 galaxies from COSMOS2020 spectro sample, z ≤ 3, ugrizy filters from HSc
  - Map the results
  - Statistically determine the “good” and the “bad” templates



BEST MODEL AT TRUE REDSHIFT

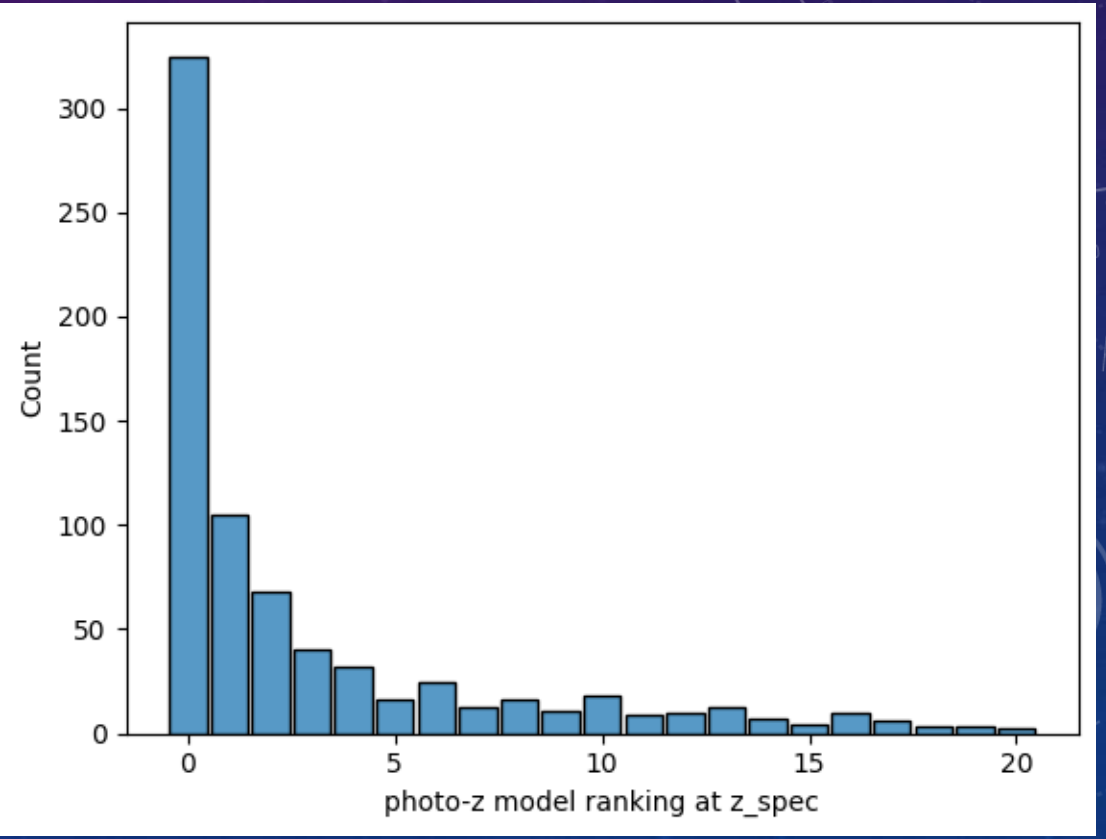
- Comparison : best model @ z_{spec} =? model at min(χ^2)
  - 1000 galaxies from COSMOS2020 spectro sample, z ≤ 3, ugrizy filters from HSc
  - Map the results
  - Statistically determine the “good” and the “bad” templates
- Focus on outliers :
  - Identify patterns of errors and corrections
  - Iterative process → automation as next step



These steps can be done quickly and efficiently with LEPHARE

IMPROVE MODEL SELECTION

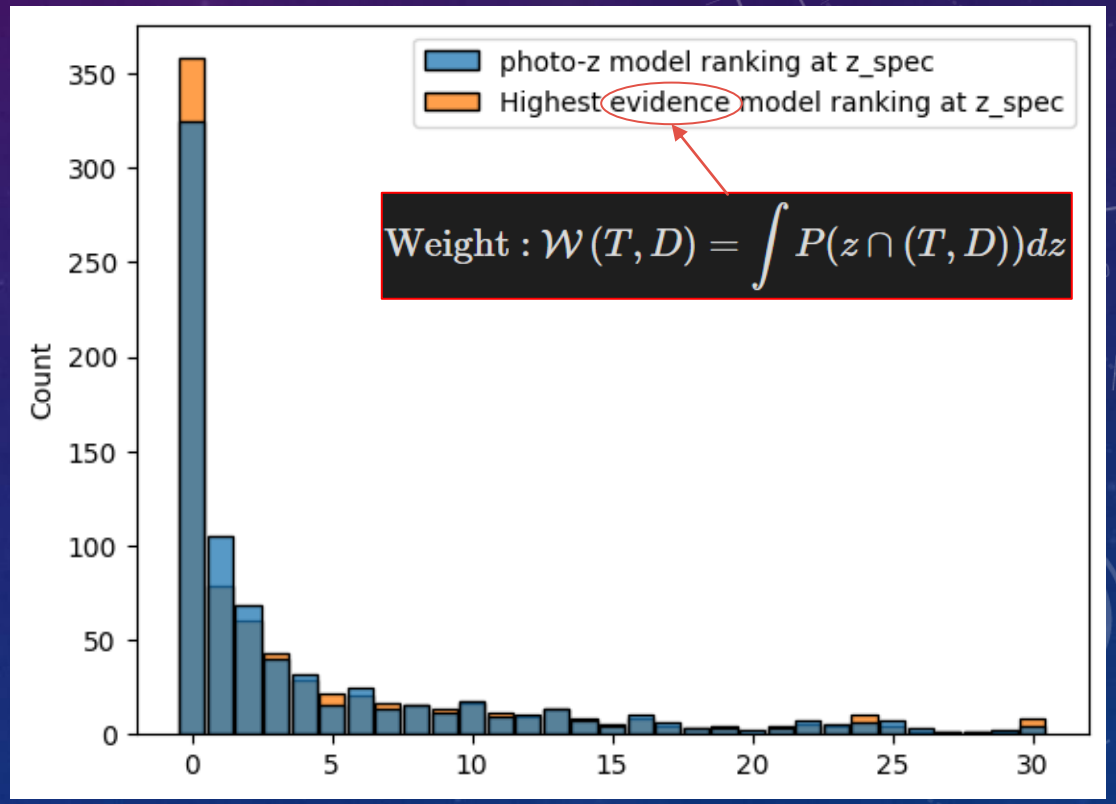
- Not a new issue : BPZ (Benitez, 2000)
- Statement : min(χ^2) does not output the best point estimate
- Idea : *rank the models according to their description of the photometry at* z_{spec}
- Only about 1/3 of perfect matches, *i.e.* when the model that yields the min(χ^2) is also the best one at z_{spec}



IMPROVE MODEL SELECTION

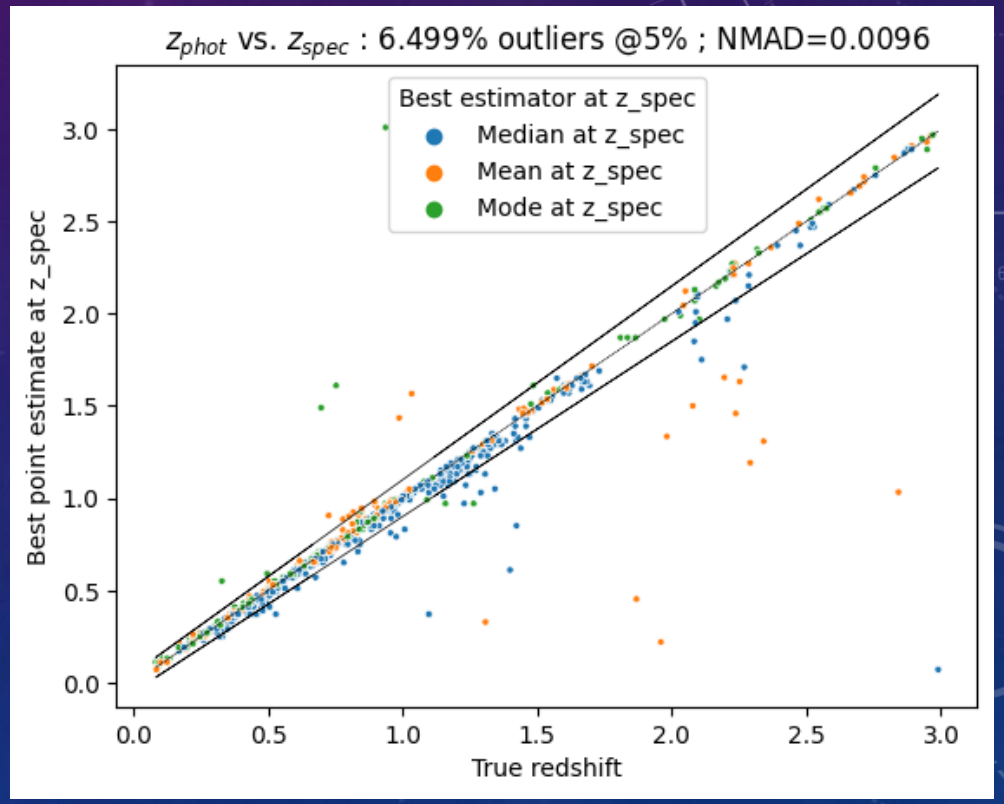
- Not a new issue : BPZ (Benitez, 2000)
- Statement : min(χ^2) does not output the best point estimate
- Idea : *rank the models according to their description of the photometry at* z_{spec}
- Only about 1/3 of perfect matches, *i.e.* when the model that yields the min(χ^2) is also the best one at z_{spec}
- Bayesian photo-z : marginalization $P(z) = \sum_{\text{models}} P(z \cap \text{model})$

$$\text{norme} = \int P(z)dz = \sum_{\text{models}} \int P(z \cap \text{model})dz$$



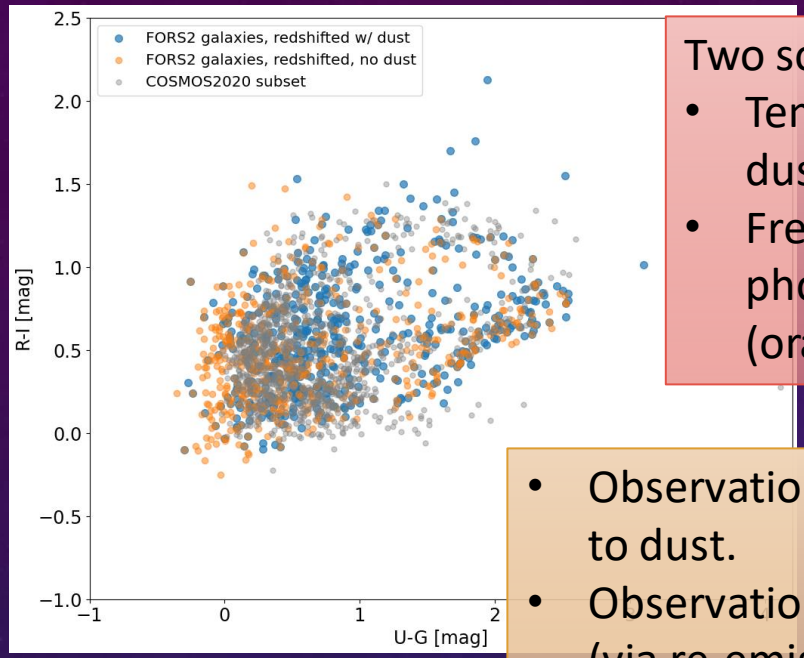
IMPROVE POINT-ESTIMATE SELECTION

- Several values can describe a PDF :
  - Mode (consistent with a min(χ^2) selection)
  - Mean (+/- σ)
  - Median (works better than the mode in LEPHARE on COSMOS2020 data)
- Which one to chose does not seem to follow a rule
- Selecting the best model and point estimate can drastically improve photo-z results



These kinds of studies can be done in our in-house photo-z tool

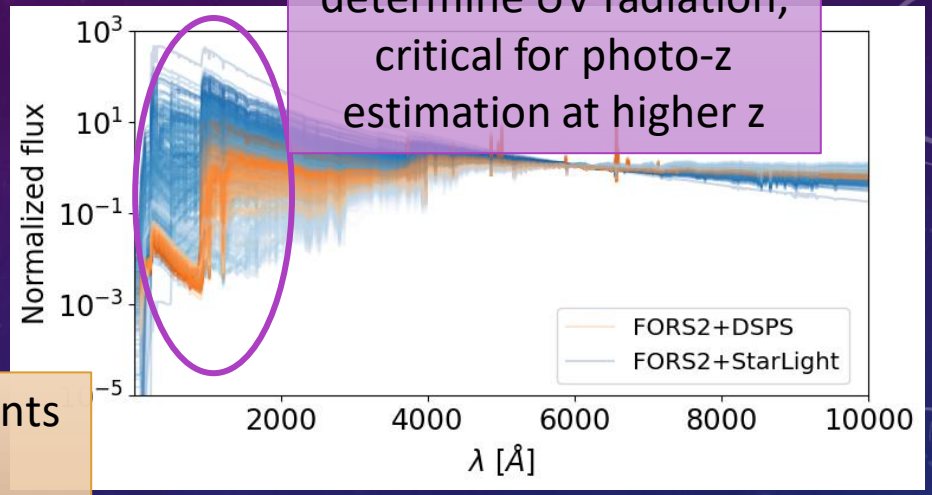
SPS, DUST AND PHOTO-Z



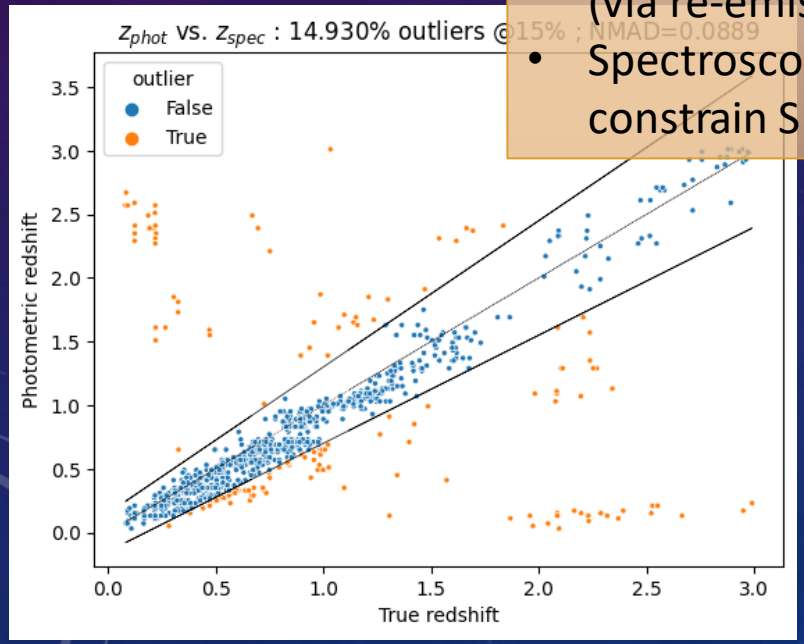
Two solutions for dust :

- Templates that include dust extinction (blue)
- Free parameter during photo-z estimation (orange)

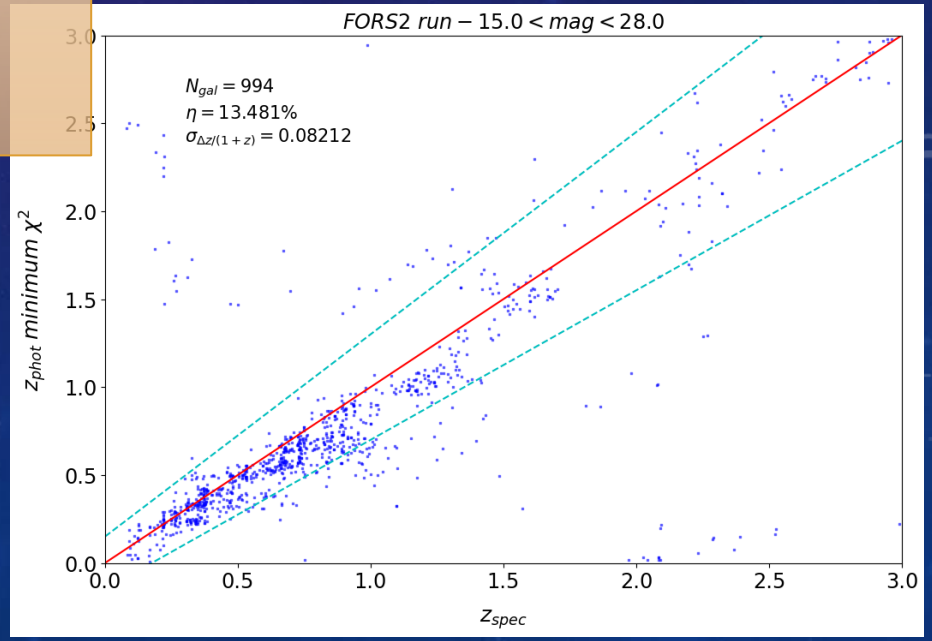
Dust extinction and SFR determine UV radiation, critical for photo-z estimation at higher z



- Observations in UV can yield SFR and hints to dust.
- Observations in infrared can constrain dust (via re-emission).
- Spectroscopy (emission lines) can constrain SFR.



FORS2 templates with dust, no extinction in photo-z estimation
← EmuLP | LEPHARE →

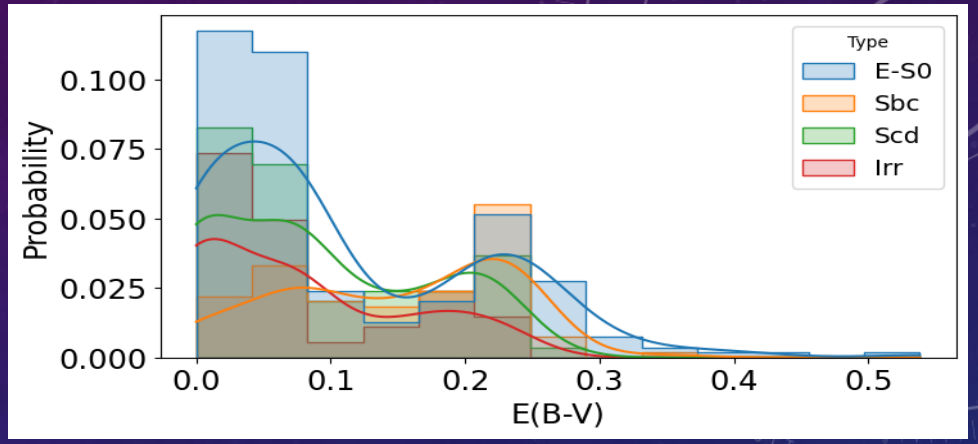


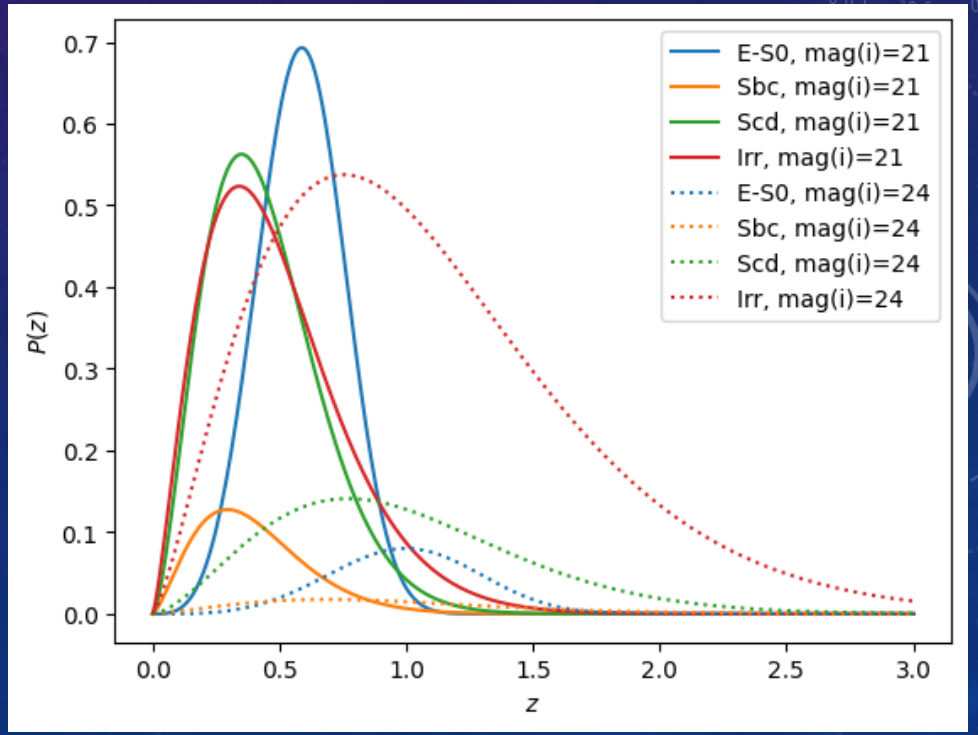
EMULATEUR LEPHARE (EmuLP)

- Template-fitting photo-z code *à la* LEPHARE (simplified) :
  - Min(χ^2)
  - Prior on n(z) based on z-VVDS
  - Added : prior on dust parameter E(B-V), derived from fit on FORS2 spectra.

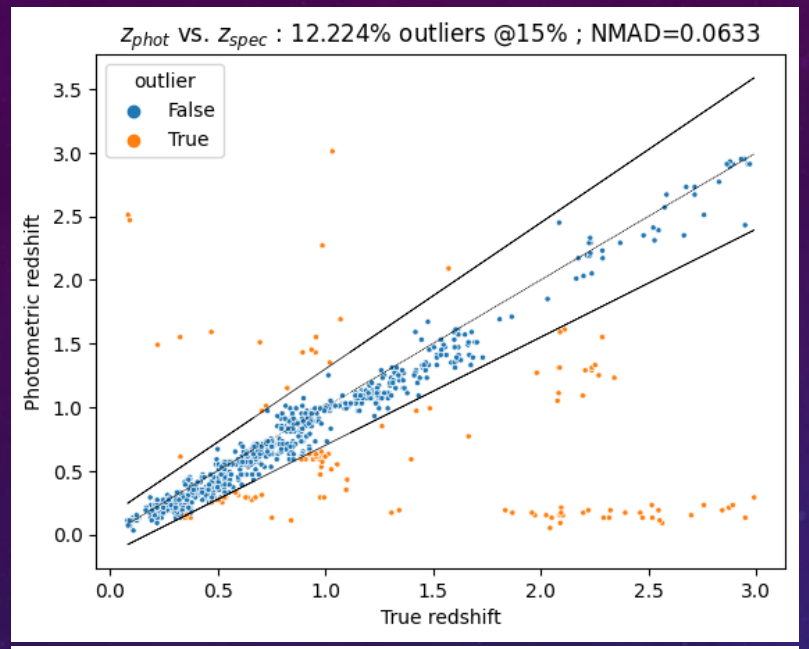
$$\chi^2 = \frac{1}{N_{filters}} \sum_{filters} \frac{(F_{obs} - F_{model})^2}{\sigma^2_{F,obs}}$$

- Written in Python (JAX for GPU compatibility), inputs in JSON
- Able to output the complete redshift probability distribution
- Includes data analysis functions (pandas)
- Platform for further tests : likelihood functions, priors, GP
- Not fast and memory-hungry : limited to small samples (~1000 galaxies) and restricted number of templates.

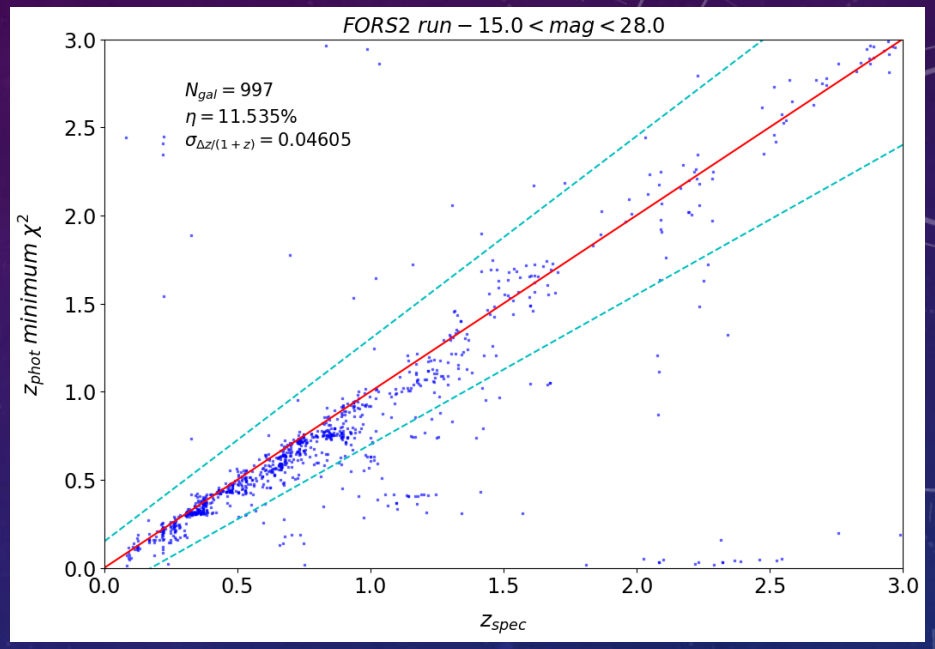


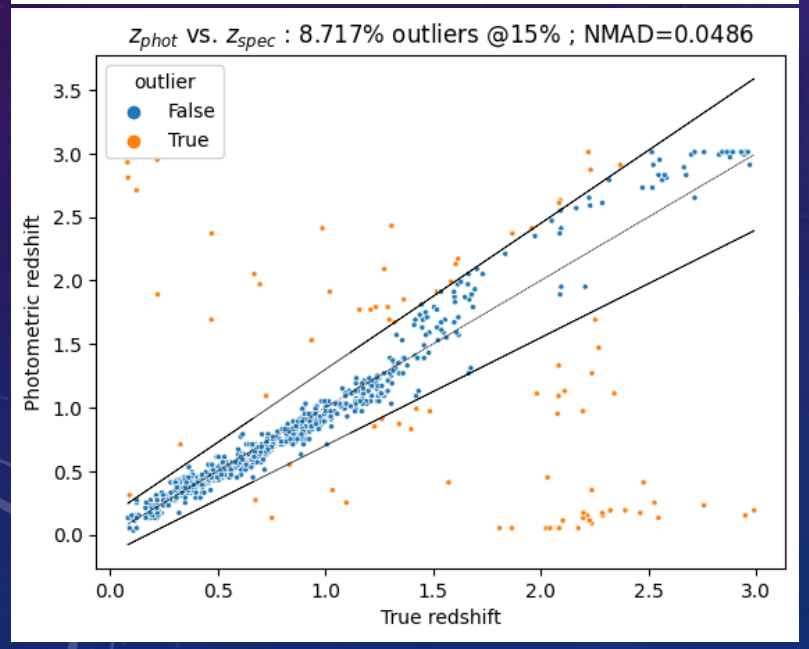


EMULATEUR LEPHARE (EmuLP)

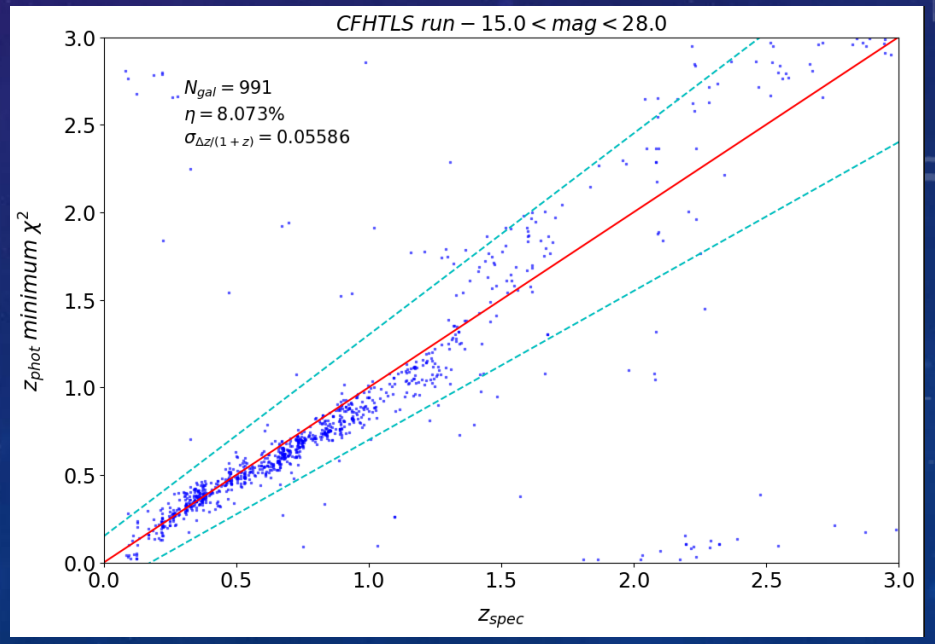


FORS2 templates
← EmuLP | LEPHARE →





CFHTLS templates
← EmuLP | LEPHARE →



TAKEAWAYS

- Aim : improve existing and/or generate new sets of templates for photo-z
- Data : FORS2, COSMOS2020, etc.
- Tools : LEPHARE++ for proper photo-z, DSPS for stellar population, *EmuLP* for in-house tests and studies
- Ideas and prospects :
  - identify spectral / stellar properties of good templates (emphasis on UV)
  - improve model selection within photo-z codes
  - explore hybrid methods (TF+ML) thanks to numerous spectra + SPS
  - ***more are welcome !***

If you have ideas, suggestions or requests related to your particular subject, please come have a chat !