Data challenges

Data Challenges

- Precisions on the DC program
    - Contents and goals of each DC
    - Precisions on the DC progression
        - We suggest to add pre-DC4 studies
    - DC organization

Data Challenges

	Meas noise	error snake	Calib noise	Color scatter	selection function	outliers	broken alpha	Pop drift dust	Lemaitre sampling	Ideal sampling	blind
DC1	green	green							green	green	
DC2	green	green			orange				green		
DC3	green	green	orange		green	orange			green		
DC4-1	green	green	green	orange	green	green			green		
DC4-2	green	green	green		green	green	orange		green		
DC4	green	green	green	green	green	green	green	orange	green		
DC5	green	green	green	green	green	green	green	green	green		orange

DC-1

- **Primary Goal :**
  - Validate the consistency of the cosmology pipeline
- **Features**
  - Only measurement noise + error snake
  - Lemaitre-like sampling & spectra
  - Ideal sampling
  - 100 realizations
- **Metrics**
  - mu - mu_true vs z (stacked over 100 realizations)
  - Coverage tests (FIM^-1 vs. empirical covmat)
- **Challenges**
  - Mostly on NaCl : make sure it can train fast enough on O(100) realizations

DC-2

- Primary Goal :
    - Validate EDRIS estimation of selection functions
    - Validate NaCl training on incomplete sample
- Features
    - Same as DC1, with incomplete samples
    - Using realistic selection functions, derived from survey data
- Metrics
    - mu - mu_true vs z (stacked over 100 realizations)
    - Coverage tests (FIM^-1 vs. empirical covmat)
- Challenges
    - Mostly on EDRIS

DC-3

- **Primary Goal :**
  - Validate propagation of calibration systematics
  - Validate non-Ia rejection in NaCl training
  - Assess importance of variable filtres (side DC-3)
- **Features**
  - Same as DC2, with calibration noise (need realistic estimate of calibration covmat by then), and non-Ia light curves.
  - Variation with variable filters (at generation)
- **Metrics**
  - mu - mu_true vs z (stacked over 100 realizations)
  - Coverage tests (FIM^-1 vs. empirical covmat)
- **Challenges**
  - Mostly on NaCl

pre-DC-4 (color scatter)

- Primary Goal :
  - Validate a method to account for the color scatter, even if only 2 bands available.
- Features
  - Same as DC3, with additional color scatter
- Metrics
  - mu - mu_true vs z (stacked over 100 realizations)
  - Coverage tests (FIM^-1 vs. empirical covmat)
- Challenges
  - Mostly on NaCl – probably a 2-color model
  - Maybe on EDRIS : how does the estimator behave if color uncertainties too large ?

pre-DC-4 (x1 / broken alpha)

- Primary Goal :
  - Integrate x1 / broken alpha standardization into EDRIS
- Features
  - Same as DC4, with broken alpha
- Metrics
  - mu - mu_true vs z (stacked over 100 realizations)
  - Coverage tests (FIM^-1 vs. empirical covmat)
- Challenges
  - Mostly on EDRIS

DC-4

- Primary Goal :
  - Integrate astrophysical systematics into generation
    - Pop drift / dust
- Features
  - Same as DC4, with more complex dag in SN generation
- Metrics
  - mu - mu_true vs z (stacked over 100 realizations)
  - Coverage tests (FIM^-1 vs. empirical covmat)
- Challenges
  - NaCl and EDRIS

Organization

- **Processing of mocks will be automated**
  - On the CC-IN2P3 farm
  - Mathieu is preparing a standard pipeline based on georges
- **Convener for each DC, who**
  - Validates the generation
  - Drives the analysis
  - Editor of the DC analysis note (mock description + DC validation analysis)
- **Conveners**
  - DC 1 : Mahmoud
  - DC 2 : Dylan ?
  - DC 3 and sqq : TBD