Université de Strasbourg

Observatoire astronomique de Strasbourg | ObAS

Debating the potential of Machine Learning in Astronomical surveys

Paris - 01/12/2023

Reionisation time field reconstruction and Model exclusion applied to WDM

Julien Hiegel

Collab. with D.Aubert, E.Thélie, R. Ibata

Hiegel et al. (2023): arXiv:2307.00609

Epoch of Reionisation: 21 cm maps

The **Epoch of Reionisation** (**EoR**) is the period where the universe transitioned from a **cold and neutral state** to a **hot and ionised state** due to the influence of the **first sources of ionising photons** (Stars, Galaxies,…)
During this epoch, the **neutral hydrogen (HI)** emits a **radio signal** (**21 cm line**) that will be observed with future radio observatories (**such as the Square Kilometre Array SKA**).

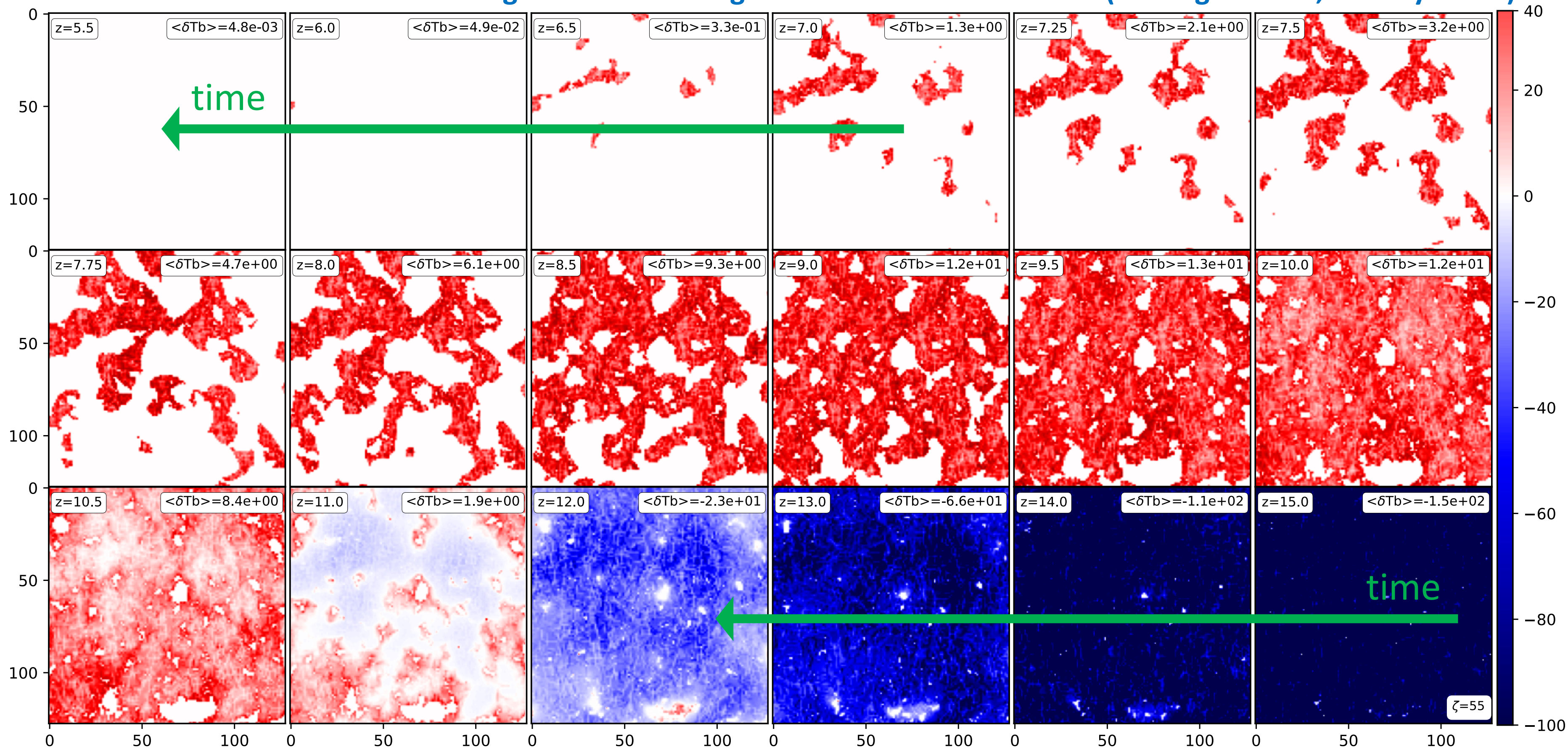
Images obtained using 21cmFAST simulation code (Mesinger+2011, Murray+2020)

Example of 21 cm maps
(pure signal)

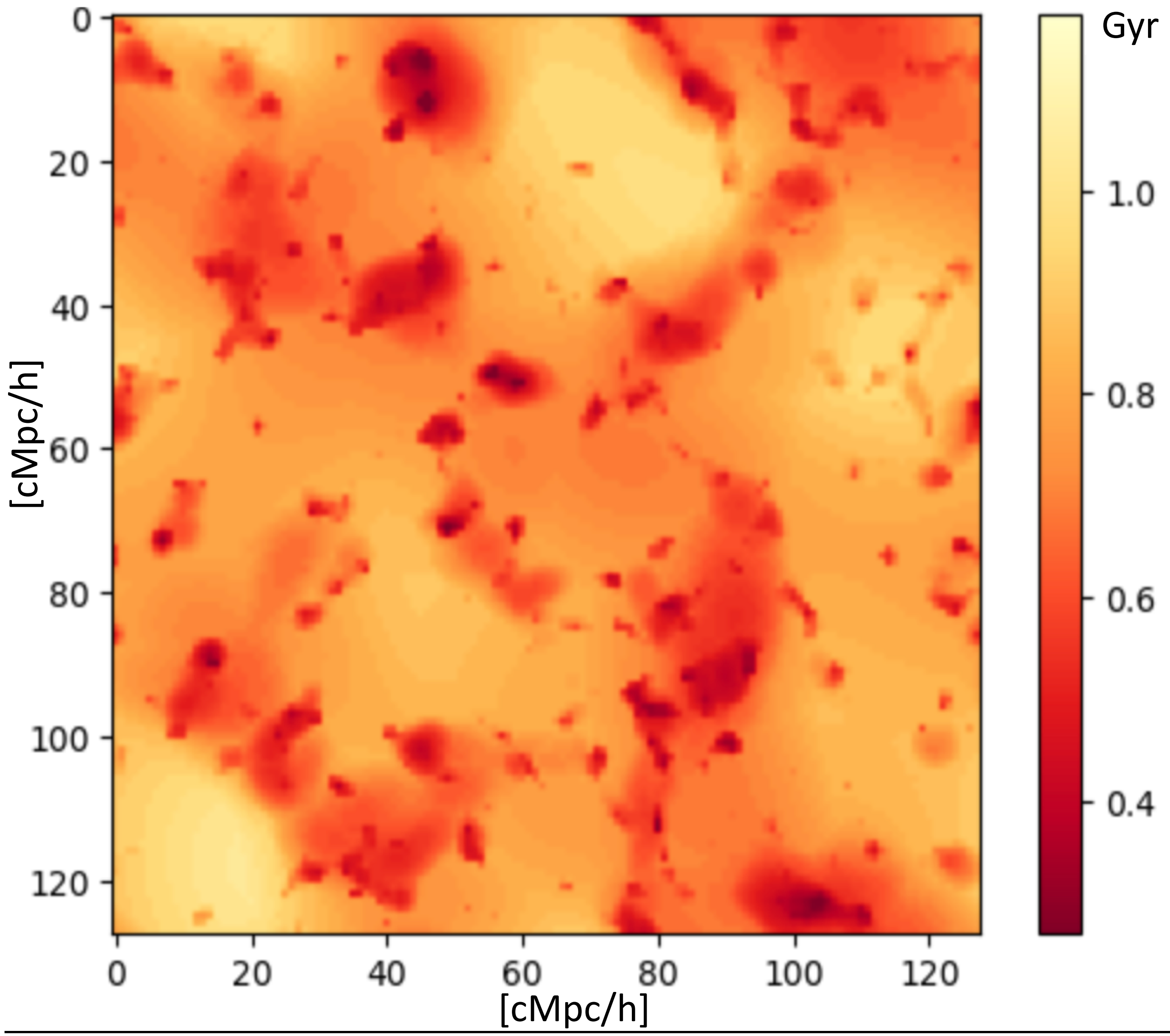
In this simulation the reionisation starts around **z=15** and ends around **z=5.5**

White regions are **Ionised bubbles of Hydrogen** that grow and merge together until all the Intergalactic medium (IGM) is reionised.

Maps size is 128x128 [**cMpc/h**]
Color bar unit is [**mK**]



Epoch of Reionisation in a single field: t_{reion}



What is the Reionisation time field t_{reion}?

It is the time at which a given region reionised where the **value of each pixel** represents the **time [Gyr] of reionisation** of this pixel

Local Reionisation process

The time of reionisation is not the same everywhere:

The **reionisation process** is **local** and **depends on the content of a given region**.

Why this field?

t_{reion} describes the whole reionisation history.

The **Reionisation seeds/Last regions to reionise** are related to the **minima/maxima** of this field (likely **densest/emptiest regions**).

We can study the **topology** of this field:

Typical size of ionised bubbles, Front speed, Abundance of reionisation seeds. **Thélie+ 2022,2023**

This is not an observable

How to recover it from observations?

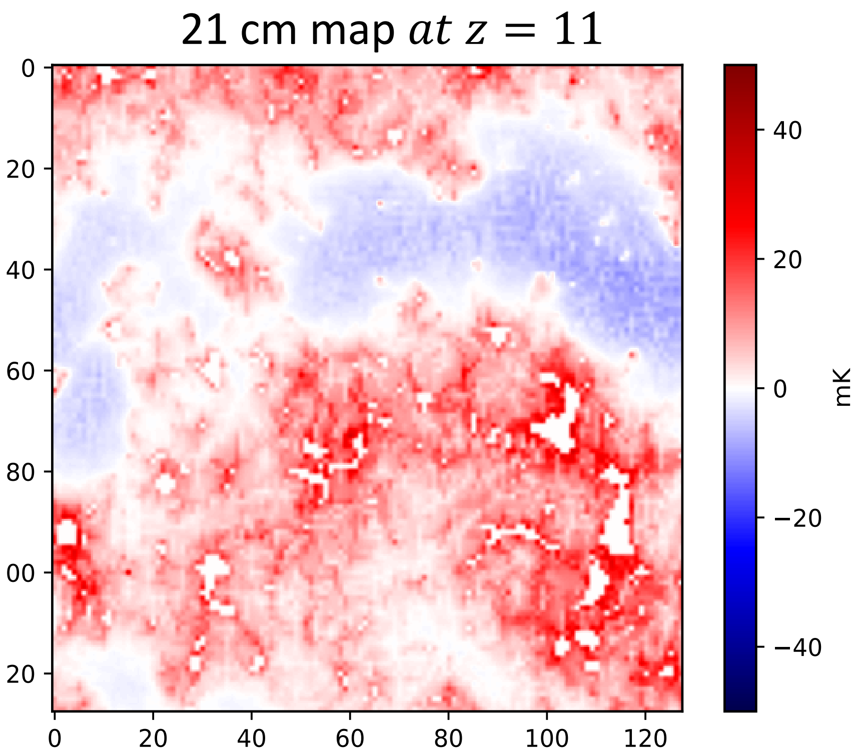
Objectives

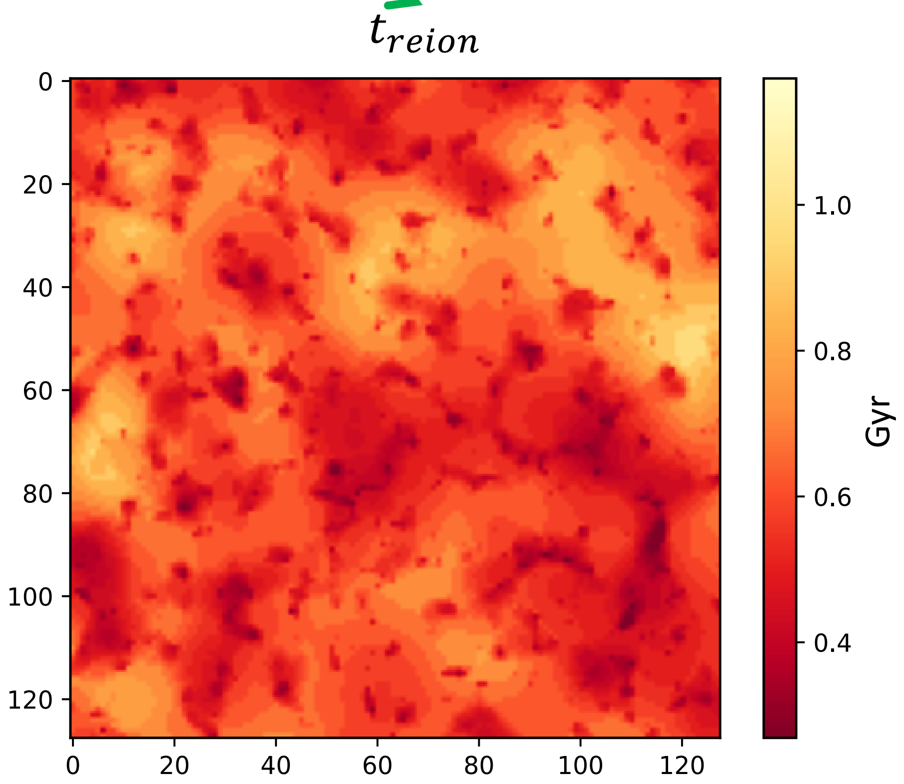
What ?

Study the EoR

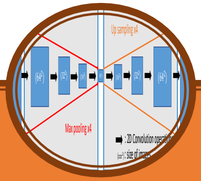
Reconstruct the time of reionisation field from 21cm signal maps and using **machine learning methods**

Reconstruction





Objectives

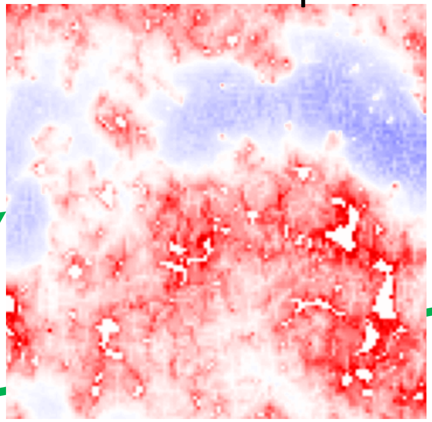


How ?

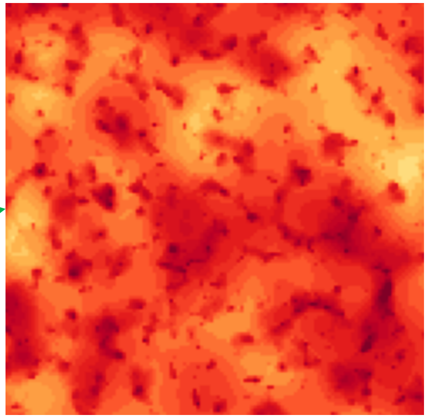
Using **21cmFAST (Mesinger+2011, Murray+2020)** simulation code to get a data set

Implementing and training a **convolutional neural network** to infer t_{reion} from 21cm maps

21 cm map



t_{reion}



Input → CNN → Output

Predictor

21cmFast simulation code & Unet Architecture

Mesinger+2011, Murray+2020

<u>In this work:</u>

- ΛCDM cosmology (**Planck+2020**)
- Size of our simulations: 256^3 pixels for 256^3 cMpch^{-1}
- Resolution set at 1 cMpch^{-1}/pixel
- **18 redshifts: $z \in [5.5, 15]$**

For a **given map**: there is **18 version** of it (1/redshift) describing its temporal evolution and these **18 maps share 1 t_{reion} map.**

<u>50 simulations</u> <u>lead to 768x50 =</u> **<u>38 400 images</u>** <u>that compose the whole data set</u>

For each field (x2)

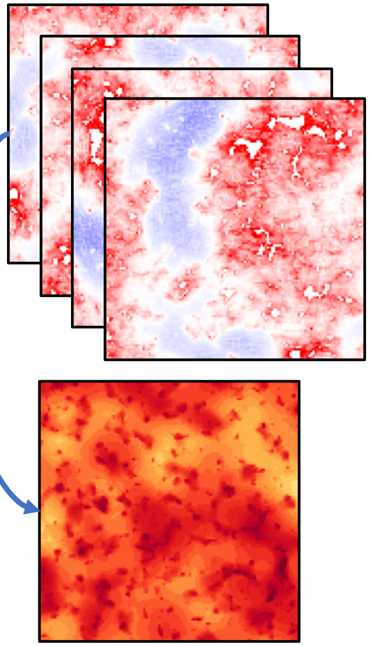
For each z (x18)

We need to divide this into **3 subsets**:

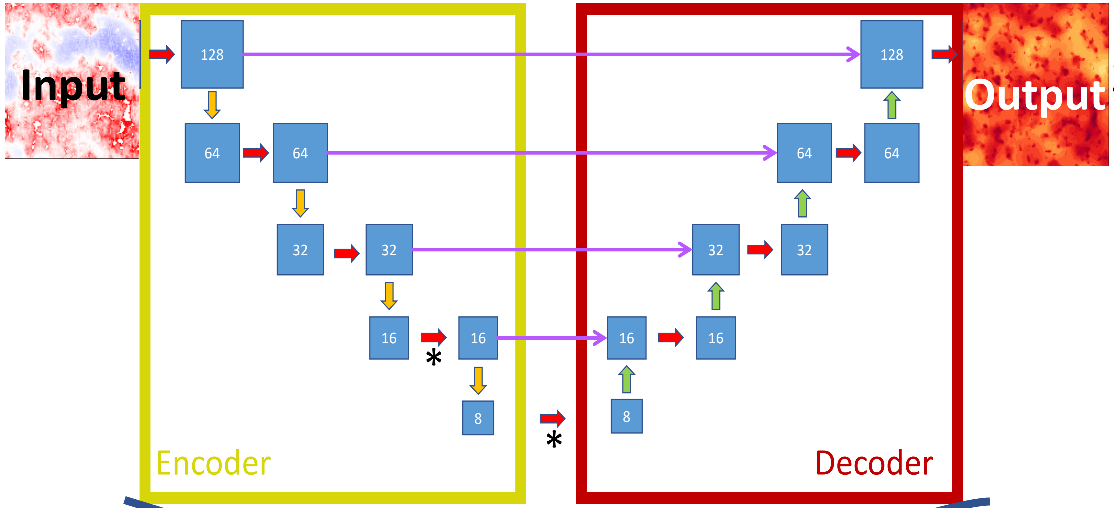
- **Training set** (Learning phase: **compute loss** and **update weights**)
- **Validation set** (Learning phase: **compute loss** only)
- **Test set** (Prediction phase: **output maps inference**)

A **trained CNN** at a **given redshift** = 1 **predictor**

18 21cm maps
For
1 t_{reion} map



31 500 maps
At a given redshift



31 500 maps

<u>1 predictor per observation redshift = 18 predictors</u>

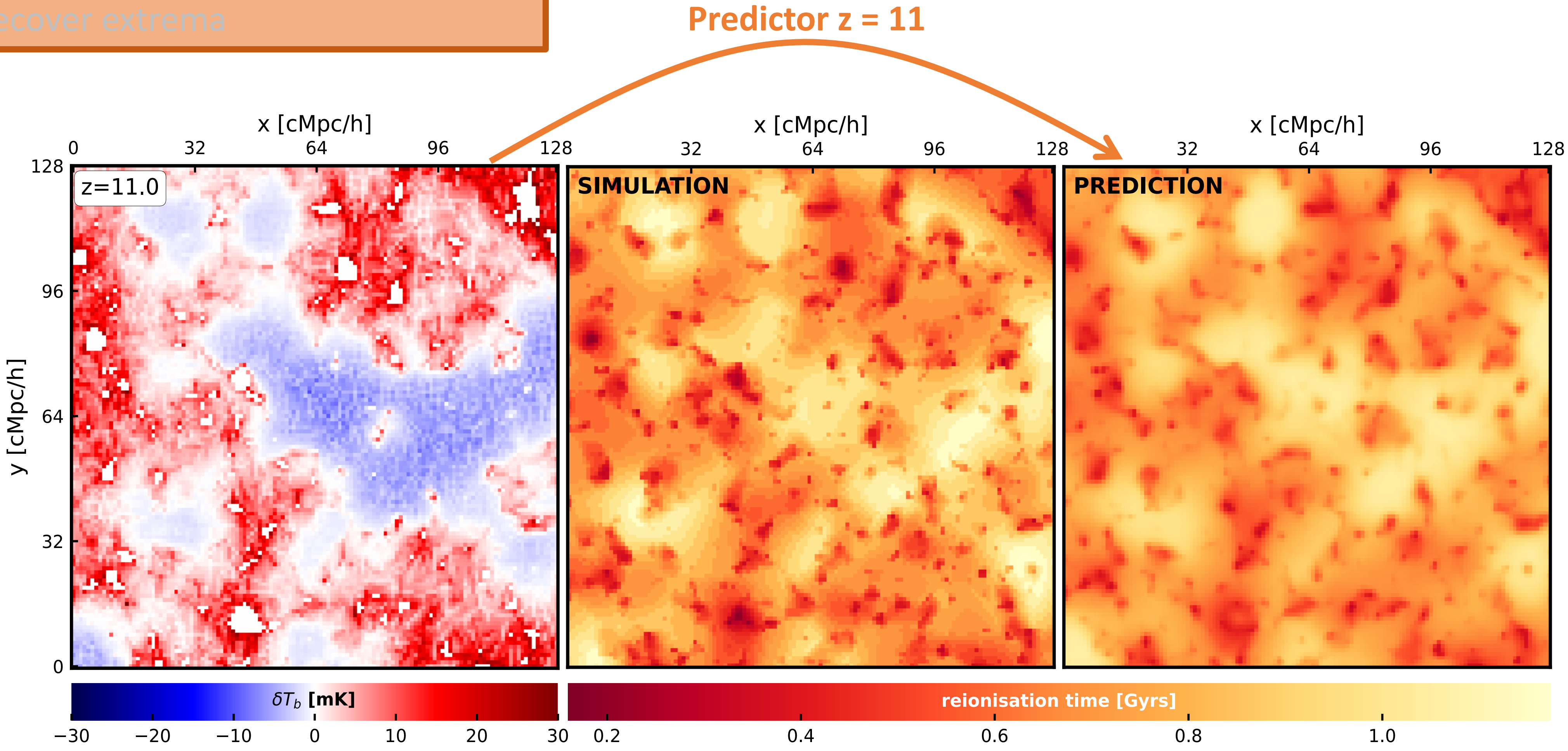
Reconstruction of reionisation time map

Example of prediction with predictor at z = 11:

- Large scale well predicted
- Blurrier/smoother version of the true field
- Difficulties to recover extrema

The 21cm map is the **current state** of HI at a given redshift:

The predictor is able to predict when a pixel in the 21 cm map will reionise in the **future** (or has reionised in the **past**)



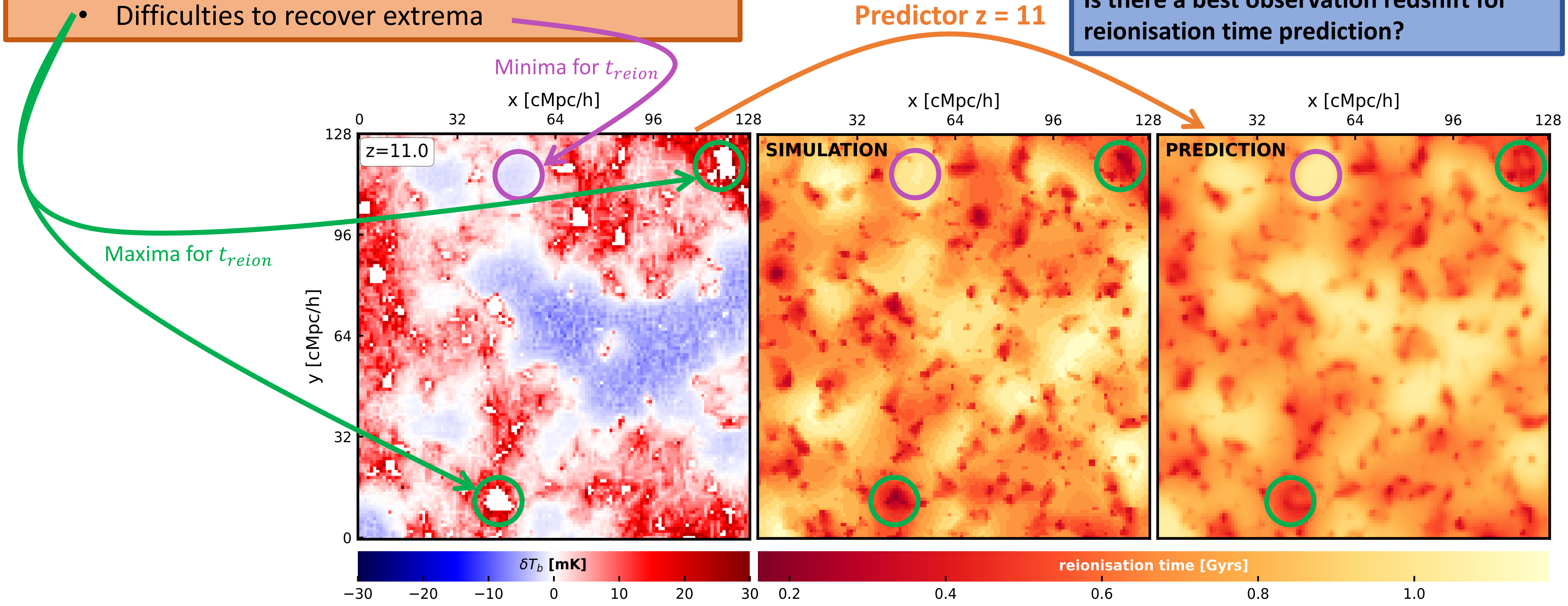
Reconstruction of reionisation time map

Example of prediction with predictor at z = 11:

- Large scale well predicted
- Blurrier/smoother version of the true field
- Difficulties to recover extrema

This operation can be made at 18 different Redshifts with 18 predictors.

Is there a best observation redshift for reionisation time prediction?



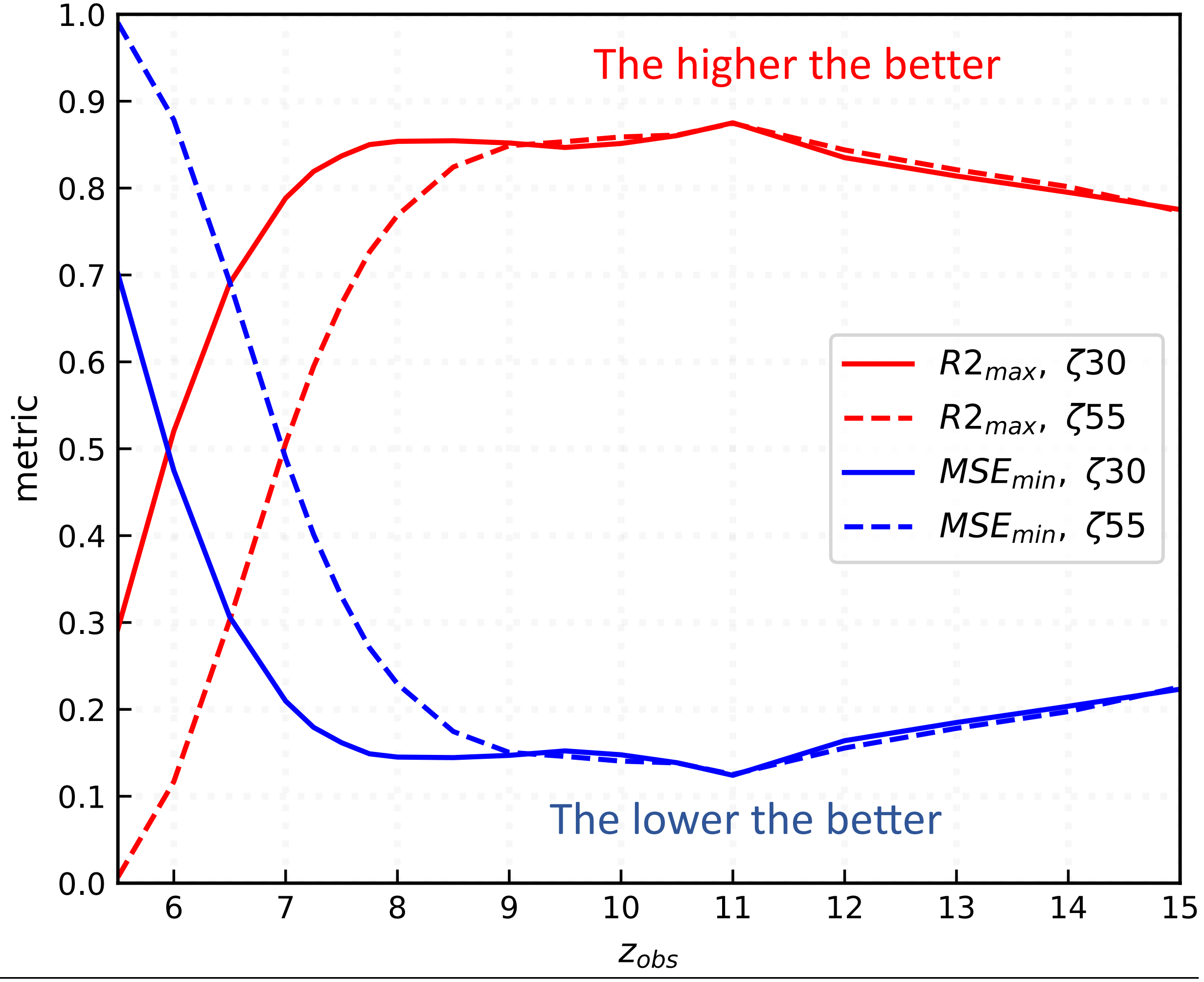
Comparison between redshifts of observation

R^2 coefficient computed on the validation set:

- **For z < 7:**
  The performances are bad:
  ⇔ Non-zero **signal fraction < 40% in 21 cm maps**
  ⇔ There is **zero signal inside ionised bubbles** that became large at this range of redshift
- **Optimal performance for z = 11:**
  In our simulations: This is the best compromise between
  ⇔ **Number of ionised bubbles that is sufficient** to locate the reionisation seeds
  ⇔ **Size of bubbles that is not too large** to get a significant amount of signal fraction
- **For z > 11**
  performances get slightly worse with increasing redshift:
  fewer ionised bubbles = We can locate fewer sources

$$R^2 = 1 - \frac{\Sigma(Pred - True)^2}{\Sigma(True - <True>)^2}$$

R^2 should be 1 for a perfect reconstruction



The performances are globally satisfying:

→ 21 cm maps contain the reionisation time information

Instrumental effects and CNN's performance

<u>New predictor for z = 8 trained with noisy 21cm maps made with SKA characteristics</u> using tools21cm (**Giri+2020)**, assuming **(Prelogovic+2022, Ghara+2016, Giri+2018**):

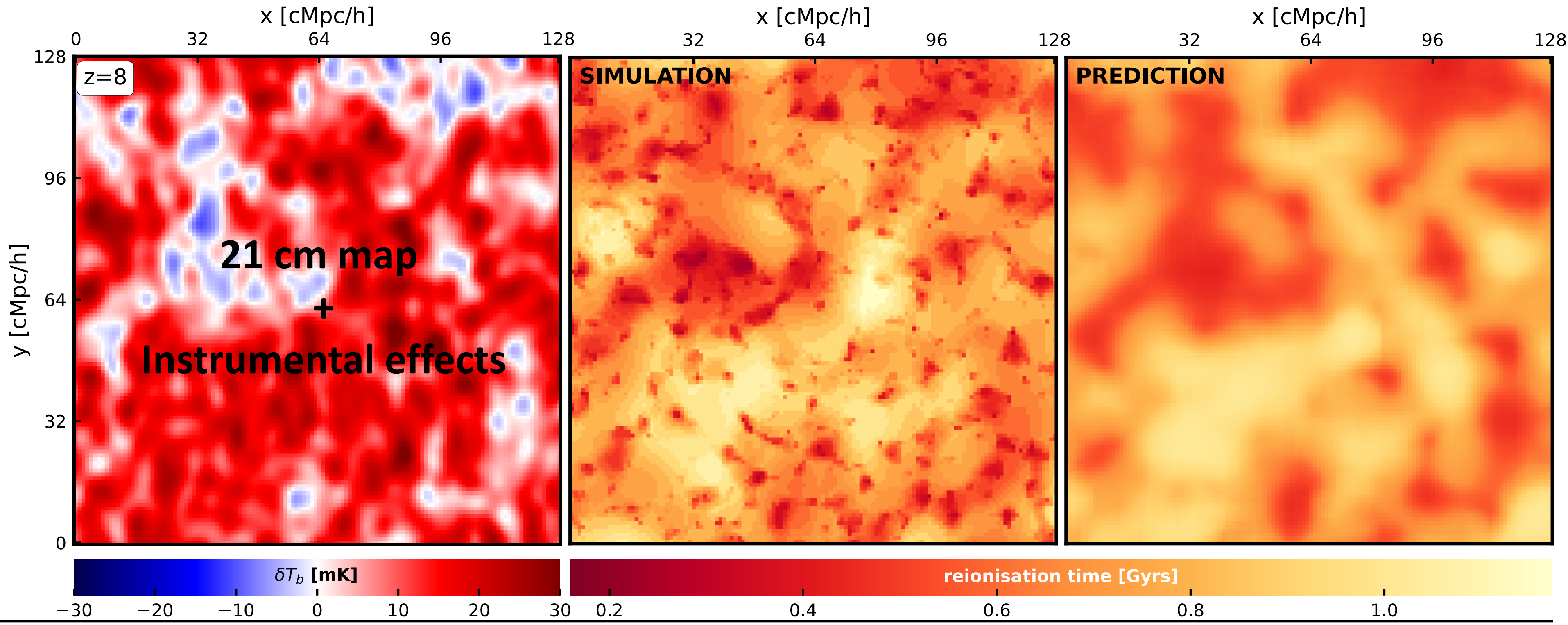
- Daily scan of 6h and 10s integration time and for a total of 1000h of observation
- A maximum baseline of 2 km
- Angular resolution of Δθ ~ 2.8 arcmin ⬄ 7.35 cMpc at this redshift

The **prediction is significantly altered** compared to the perfect observation scenario:

- Large-scale structures are less prominent, appearing buried or obscured and a lot blurrier

<u>z = 8</u>:

Best compromise between reconstruction performance and noise



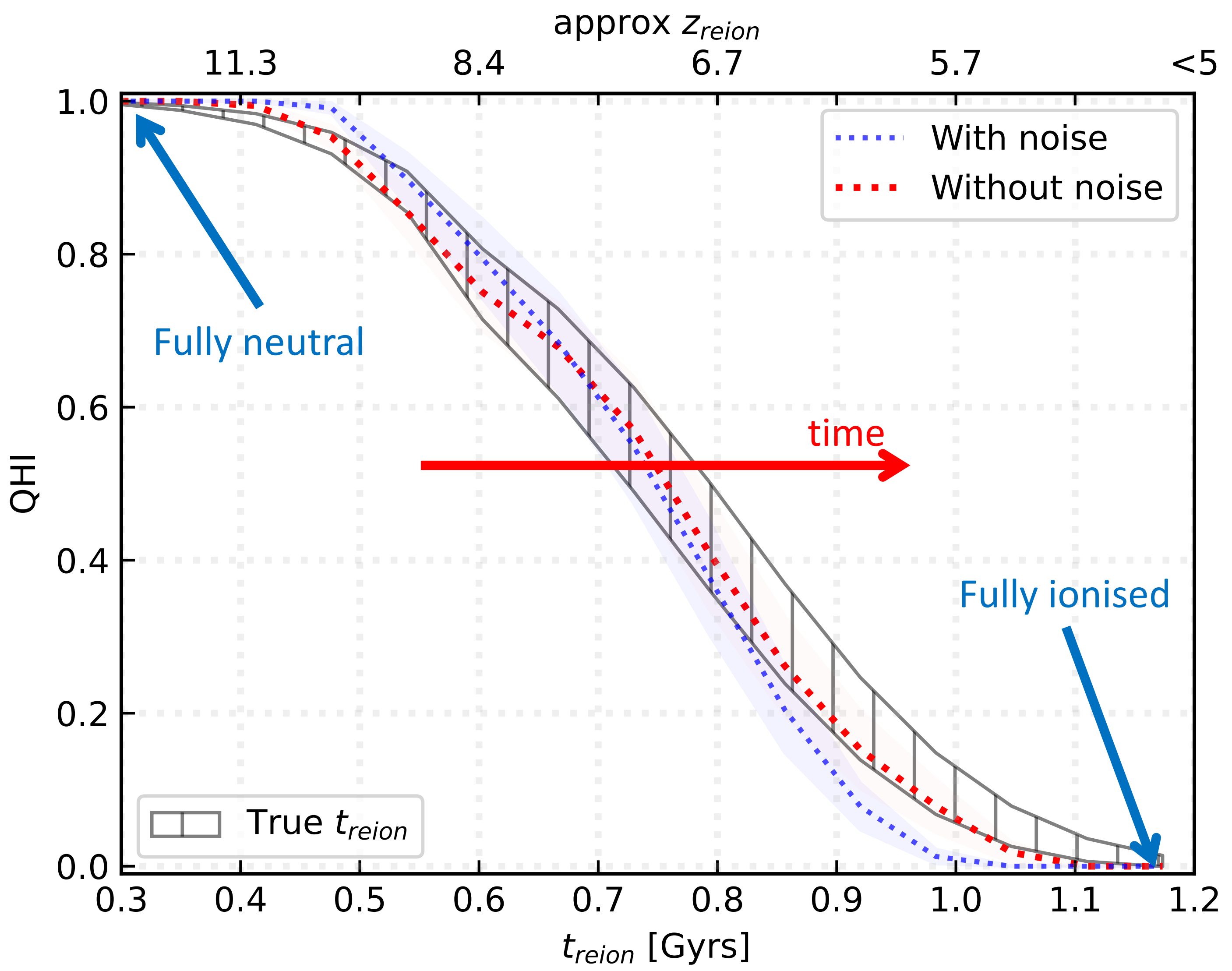
Instrumental effects and CNN's performance
fraction of neutral Hydrogen

t_{reion} allows to reconstruct the average neutral fraction history $Q_{HI}(t)$

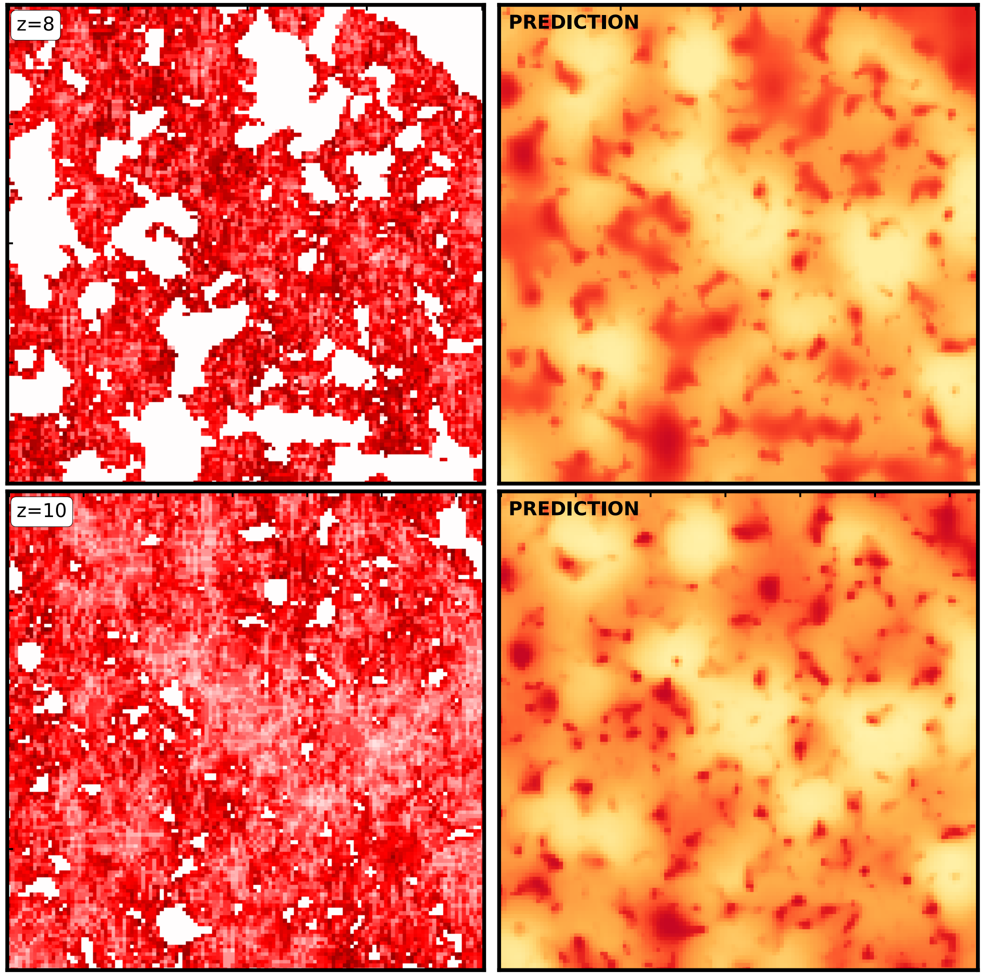
<u>For 21 cm maps without noise</u>: (red)
The predictor gives a **satisfying reionisation history**

<u>For noisy observations</u>: (blue)
The predictor **starts** the **reionisation later** and tends to **finish it earlier** than true field.
-> It suffers from the smoothing since the **extrema are erased**.

Nevertheless the reconstruction of the global reionisation history is satisfying



Questions



What happens when **some data is given to a "wrong" predictor (with wrong underlying model)**??

-> Ideally data given to a predictor with the right model should give the **same t_{reion} map no matter the redshift of observation**

→ **We may compare results at 2 redshift of observation**

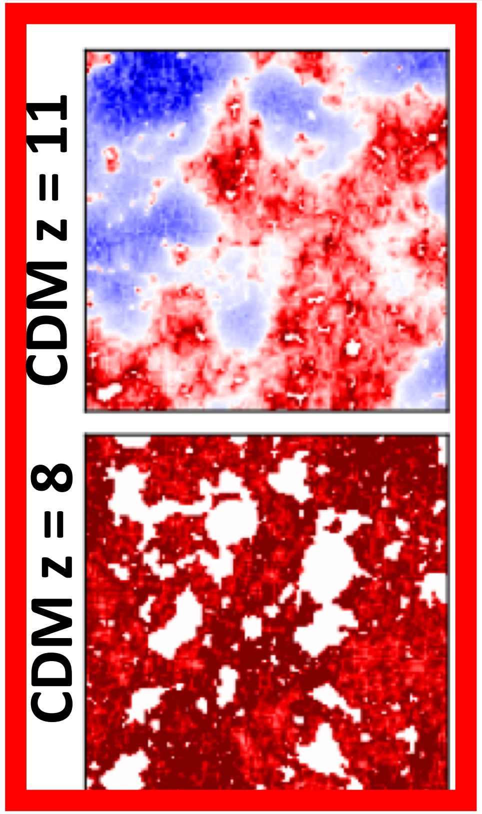
If we notice **an inconsistency between 2 reconstructions at 2 different observation redshifts,**
Can we then use it to **exclude predictors** ??

→ A **model** is **linked** to **each predictor**: **excluding a predictor means excluding a model**

We test this method to exclude WDM models (work in progress)

CDM Model VS Predictors

We arbitrarily set one model as **reference**: **CDM** is in this study our **mock observations.**



CDM Model VS Predictors

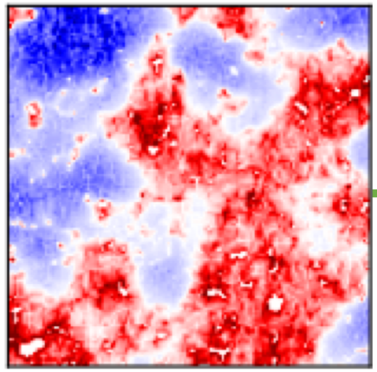
We arbitrarily set one model as **reference**: **CDM** is in this study our **mock observations.**

By training CNN with **DM models (3keV, 5keV, 7keV, CDM) and z = 8, 11**: We get a list of **predictors** that have their **own model**.

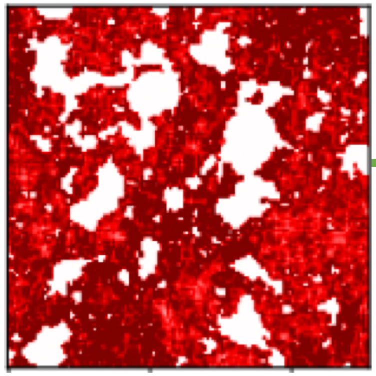
Identical CDM model as previously
WDM models have a **similar reionisation scenario**

Example: We have 2 Predictors for the model WDM 3 keV
1 for each redshift (8 and 11)

CDM z = 11



CDM z = 8



Predictor z = 11

- CDM
- WDM 3keV
- WDM 5keV
- WDM 7keV

Predictor z = 8

- CDM
- WDM 3keV
- WDM 5keV
- WDM 7keV

Lower limit on the thermal relic WDM particle mass:

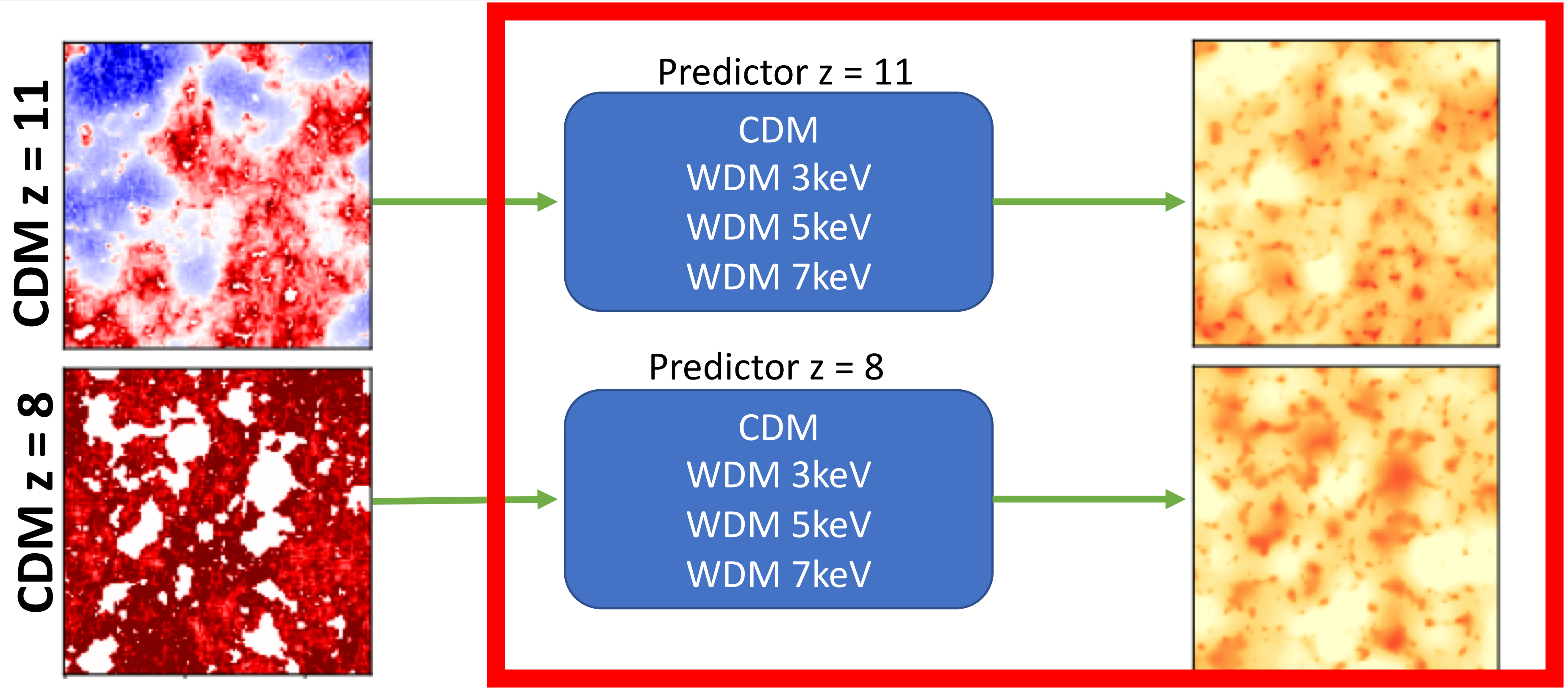
Iršič+2023: **5.7 keV** (Ly α)
Enzi+2021: **6.048 keV** (Strong Grav. lensing, Ly α, MW satellites)
Nadler+2020a: **6.5 keV** (MW satellites)

CDM Model VS Predictors

We arbitrarily set one model as **reference**: **CDM** is in this study our **mock observations.**

By training CNN with **DM models (3keV, 5keV, 7keV, CDM) and z = 8, 11**: We get a list of **predictors** that have their **own model**.

Thanks to the mock observations (**CDM**), we **infer t_{reion} with <u>each predictor</u>** (For the 4 DM models and for the 2 redshifts).



CDM Model VS Predictors

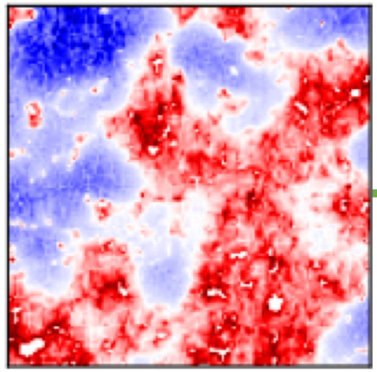
We arbitrarily set one model as **reference**: **CDM** is in this study our **mock observations.**

By training CNN with **DM models (3keV, 5keV, 7keV, CDM) and z = 8, 11**: We get a list of **predictors** that have their **own model**.

Thanks to the mock observations (**CDM**), we **infer t_{reion} with each predictor** (For the 4 DM models and for the 2 redshifts).

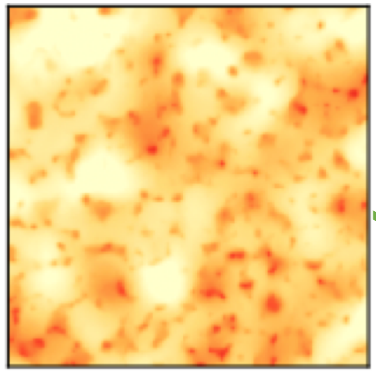
We compare the predictions and **estimate** whether the **predictor fits the observation** -> **Model Exclusion.**

CDM z = 11

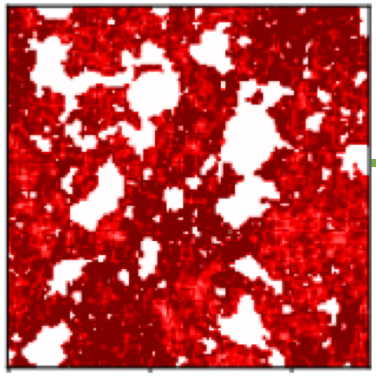


Predictor z = 11

CDM
WDM 3keV
WDM 5keV
WDM 7keV

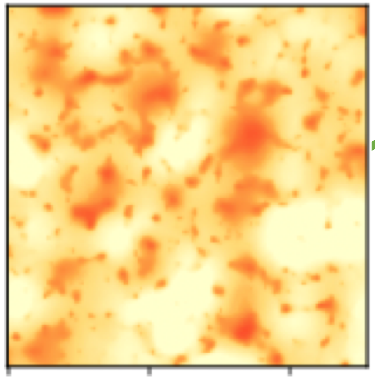


CDM z = 8



Predictor z = 8

CDM
WDM 3keV
WDM 5keV
WDM 7keV



For a predictor matched with the observation model, t_{reion} from 2 z_{obs} should be (statistically) identical

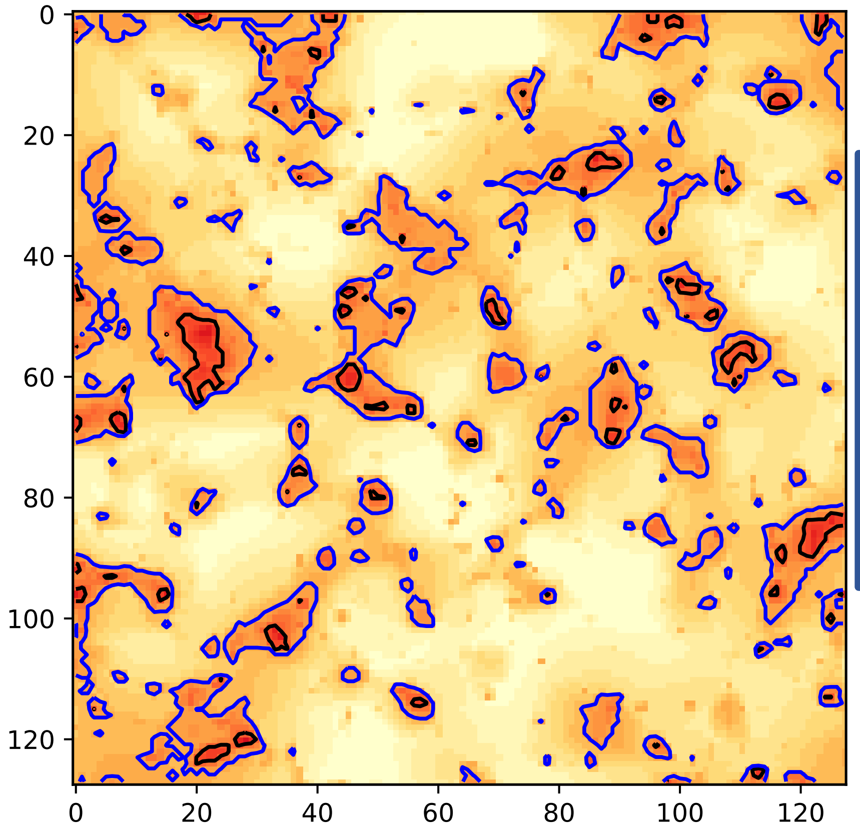
Comparator

Autoconsistency check with t_{reion} isocontours

Reionisation Time Field isocontours taken at t = [0.4, 0.6] Gyr

Depicts the **interface between ionised and neutral regions/bubbles at a given time**

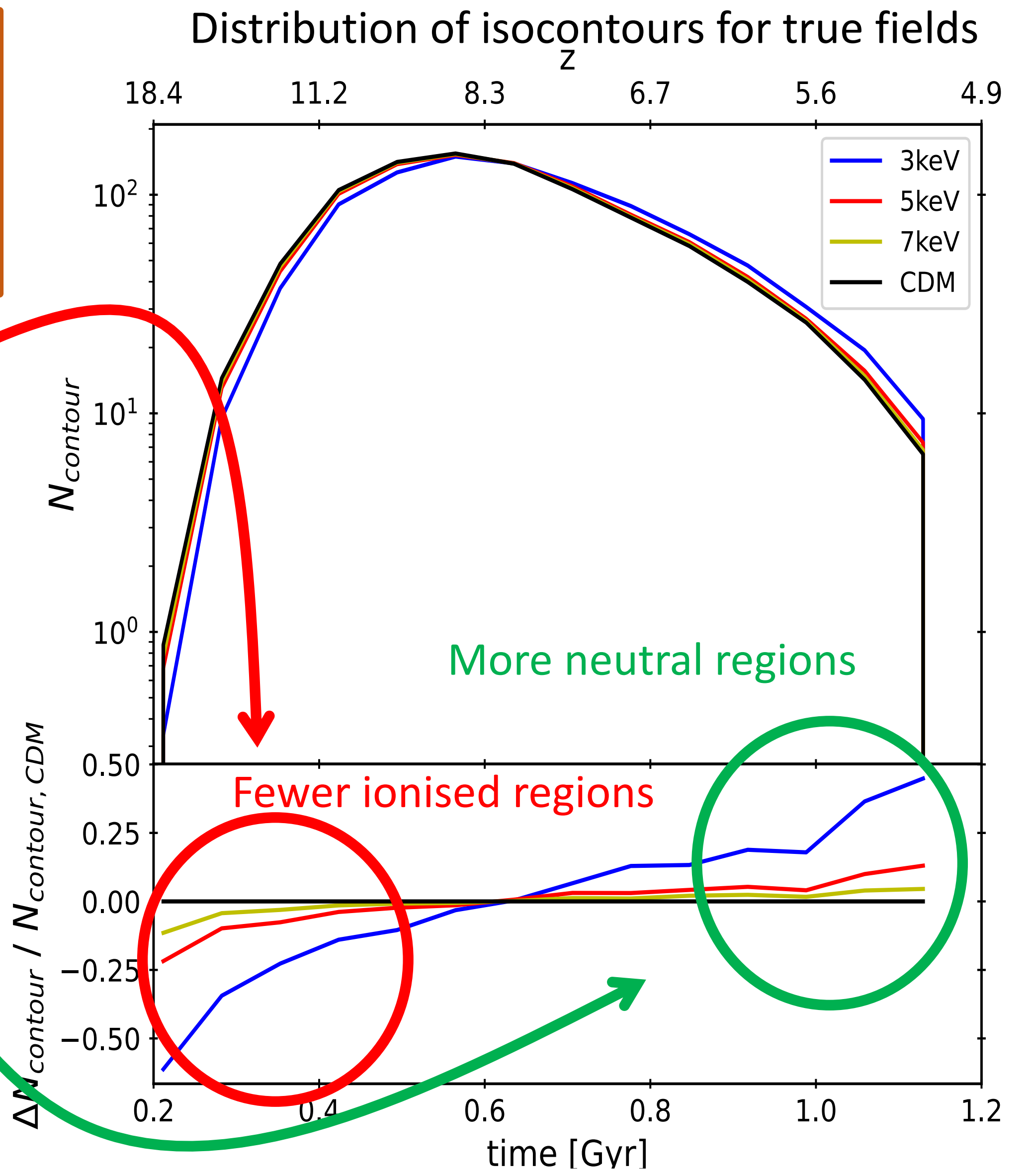
By measuring the contours, we extract information on the distribution and the size of ionised/neutral bubbles (***Thélie+ 2023)***



WDM models have **fewer contours at early times** and **more contours at late times** than CDM

-> Fewer **ionised regions** at early times

-> More **neutral regions** at the end of reionisation



Autoconsistency check with t_{reion} isocontours total Length L

Work In progress (Hiegel+ in prep)

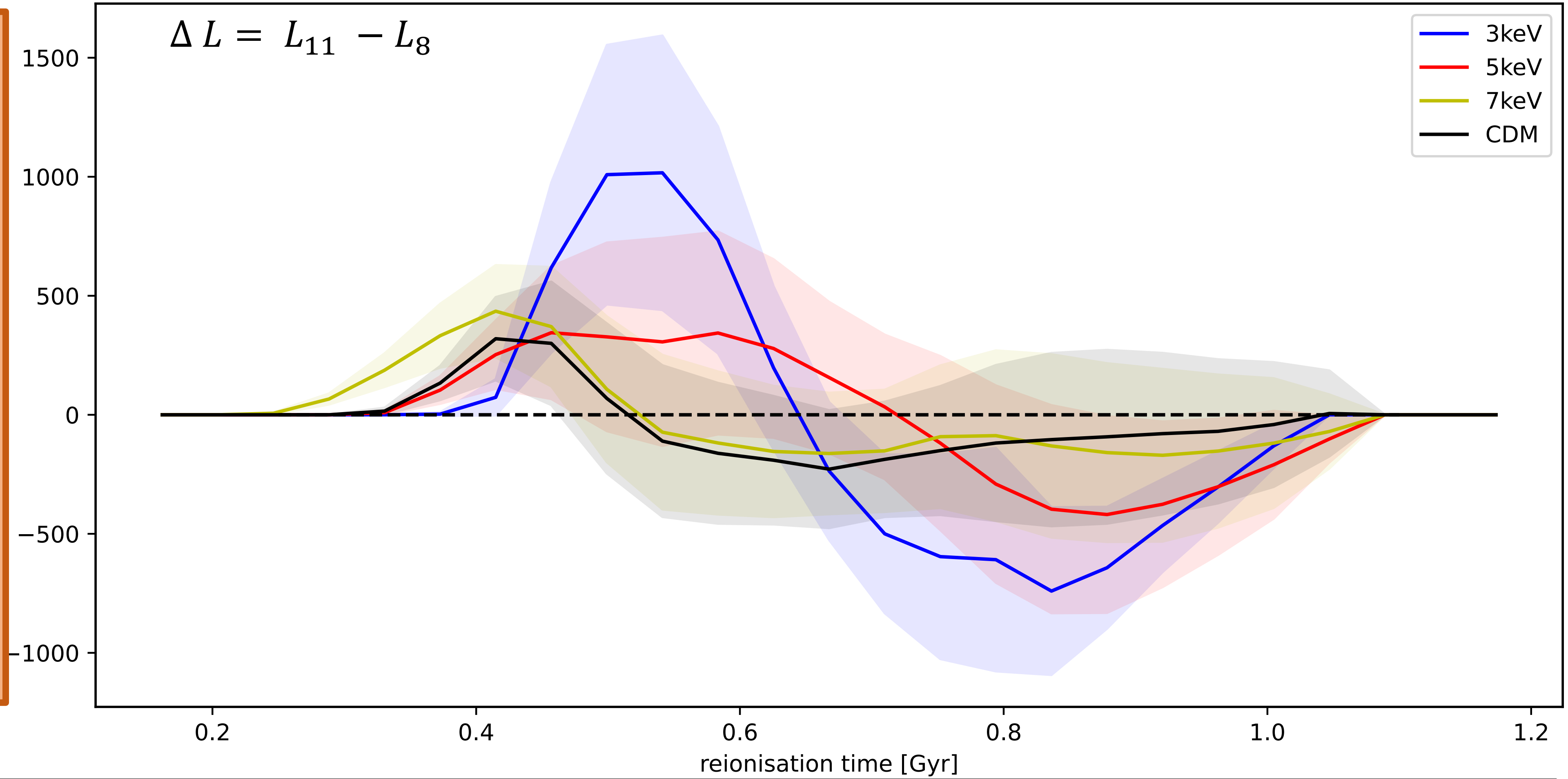
For a given predictor, as the **predictions should be statistically identical** we should have the **same isocontours for z=11 and z=8**

Then we compare the **isocontours total length** from z_{obs}=11 and z_{obs} =8 to test the auto consistency of a model: **Here even the CDM isn't perfectly auto consistent**

<u>7keV Predictor</u>
Has the same level of auto consistency as **CDM**

<u>5keV & 3keV Predictors</u>
Seem to present a larger level of inconsistency between isocontours at z=11 and 8: we can possibly exclude these predictors on the basis of this statistic

We are working on **other methods** or **metrics** to be able **to exclude models** in a more systematic way



Conclusions

Reionisation time recovery: arXiv:2307.00609

The **reionisation time field information is encoded in the 21 cm signal** and we can reconstruct t_{reion} from 21 cm maps

The **predictor performance depends on z_{obs}** for a given reionisation scenario and there is an optimal z_{obs}. We get **good reconstruction at large scales**, however there is **small scales issues**

Predictions with **noisy observation** are in broad agreement with ground truth at large scales but the **smallest scales are totally missing** on the prediction

Model Exclusion applied to WDM models: Work in progress

We can exclude models of WDM with 2 observations but **it requires more investigations to get more quantitative results**

Perspectives

How to recover smallest scales and increase general performance ??

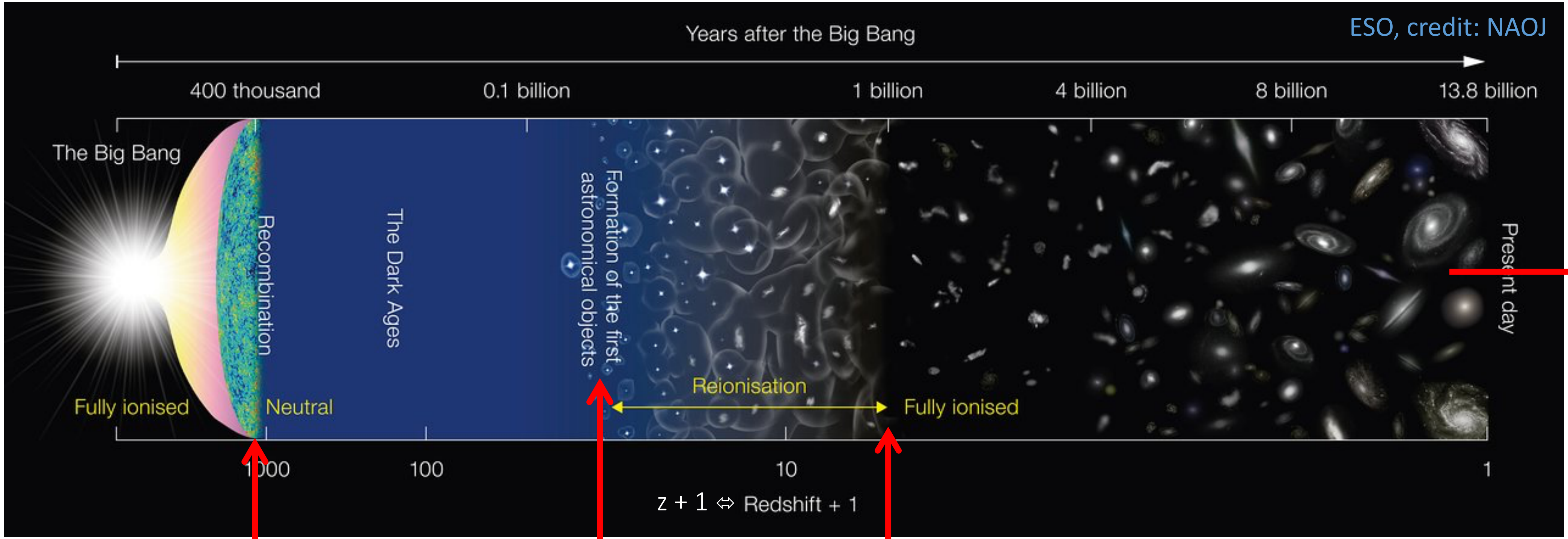
Generative Adversarial Networks (GANs) might be a solution (Ullmo+ 2020)
Implementing Attention Block to improve feature extraction can boost the performance (Chen+ 2016)

How can we Validate/exclude the underlying model of a predictor ??

What metric (e.g**. P(k), Npeak**) can we use to exclude models ?
Can we give a **score** that tells the **confidence** we have in a given predictor

Is this exclusion process viable with noise?

Overview of the Universe history



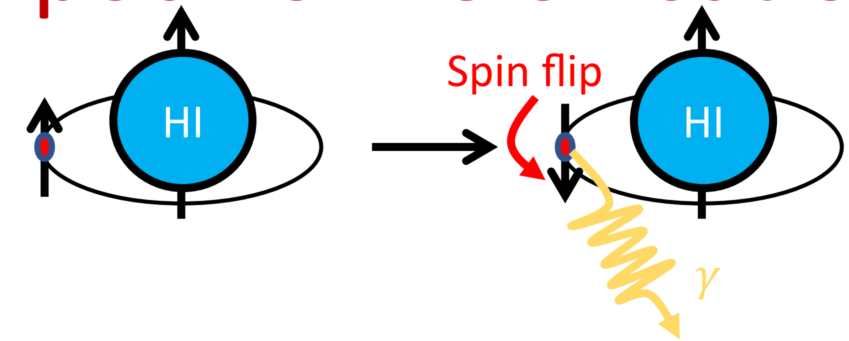
Future

Cosmic Dawn

First emitted light

End of Epoch of Reionisation (EoR) :
last big transition of the universe

Epoch of Reionisation: 21 cm maps



Emitted Photon:
f = 1 420 MHz
λ = 21 cm

➔ Will be redshifted at lower Frequency where it can be observed in the range **50-350MHz** with the **Square Kilometre Array (SKA, Braun+ 2019)**

Images obtained using 21cmFAST simulation code (Mesinger+2011, Murray+2020)

Example of 21 cm maps

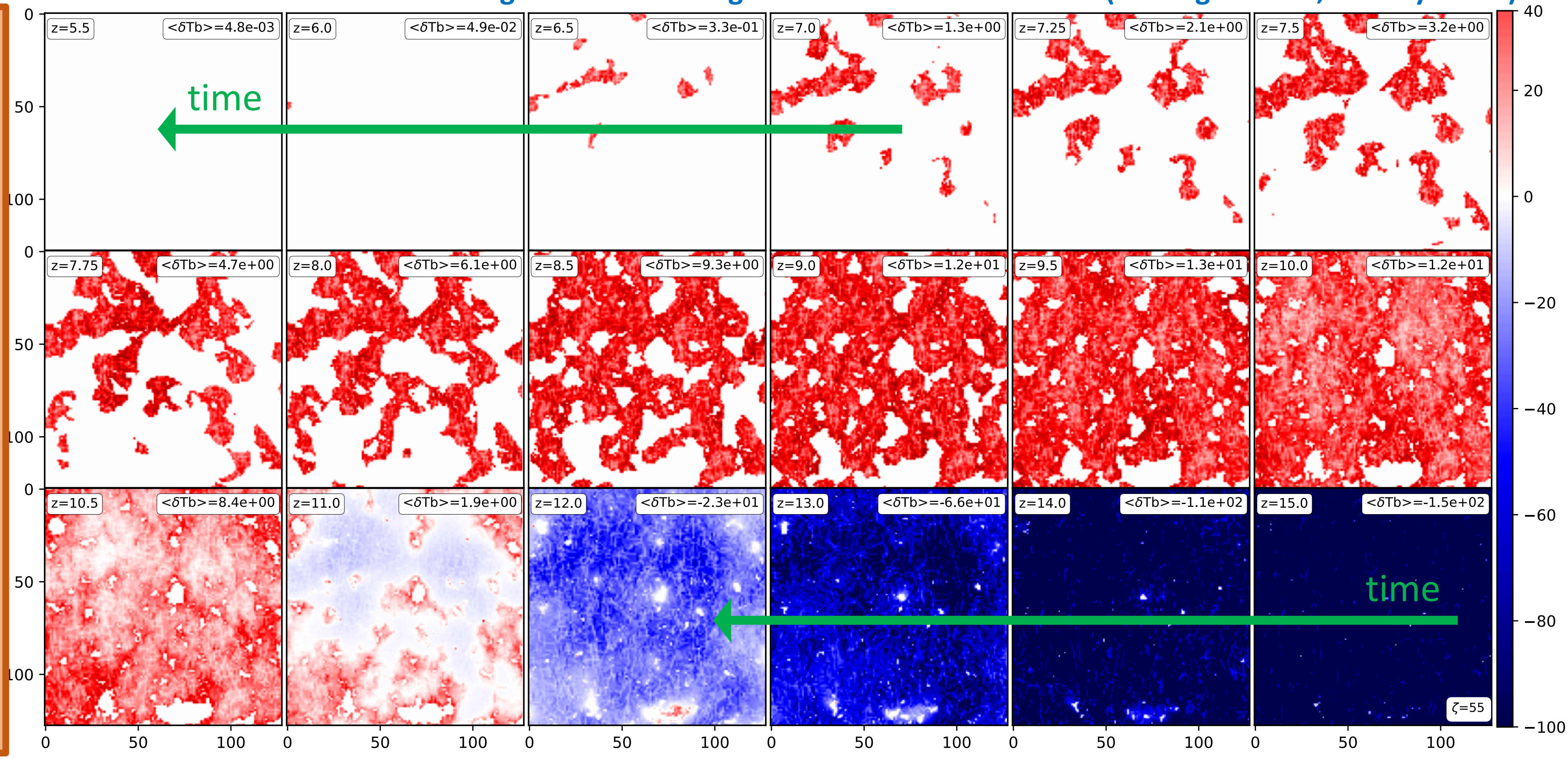
In this simulation the reionisation starts around **z=15** and ends around **z=5.5**

Blue regions (negative values) means the T_{HI} is colder than T_{CMB} and the 21 cm line is seen in **absorption**

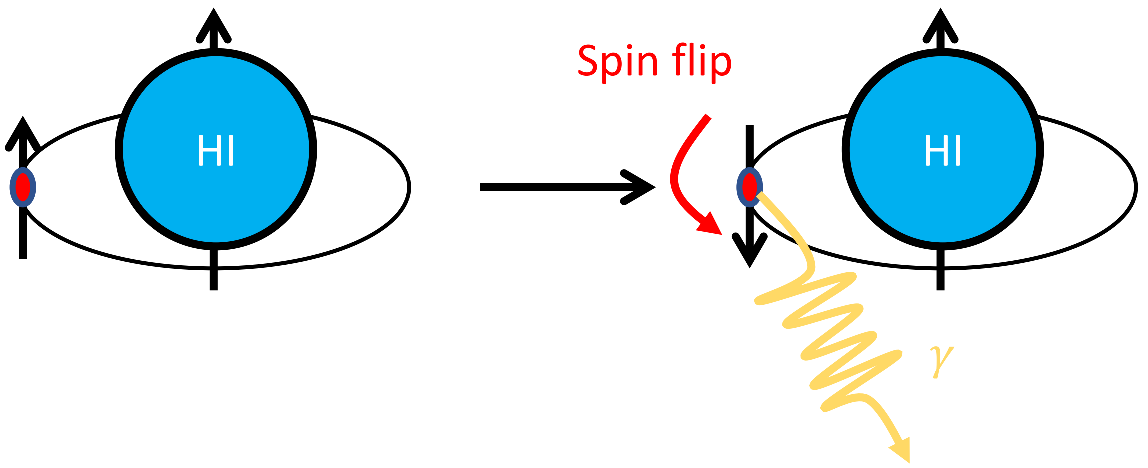
Red regions (positive values) means the T_{HI} is warmer than T_{CMB} and the 21 cm line is seen in **emission**

White regions are **Ionised bubbles of Hydrogen** that grow and merge together until all the Intergalactic medium (IGM) is reionised.

Maps size is 128x128 [**cMpc/h**]
Color bar unit is [**mK**]



21cm signal and Square Kilometer Array (SKA)



Emitted Photon:
f = 1 420 MHz
λ = 21 cm

Will be redshifted
at
lower Frequency

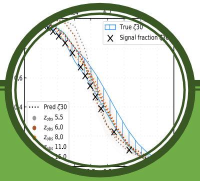
SKA-low antennas – Credit: https://www.skao.int/en/explore/telescopes/ska-low

SKA-low (late 2020s):

130k+ Antennas in Australia
0.4 km^2 collecting area
Frequency range: 50-350MHz

It is the necessary range to observe the 21cm line through the EoR

Objectives



Why ?

A single map describes the whole history of reionisation

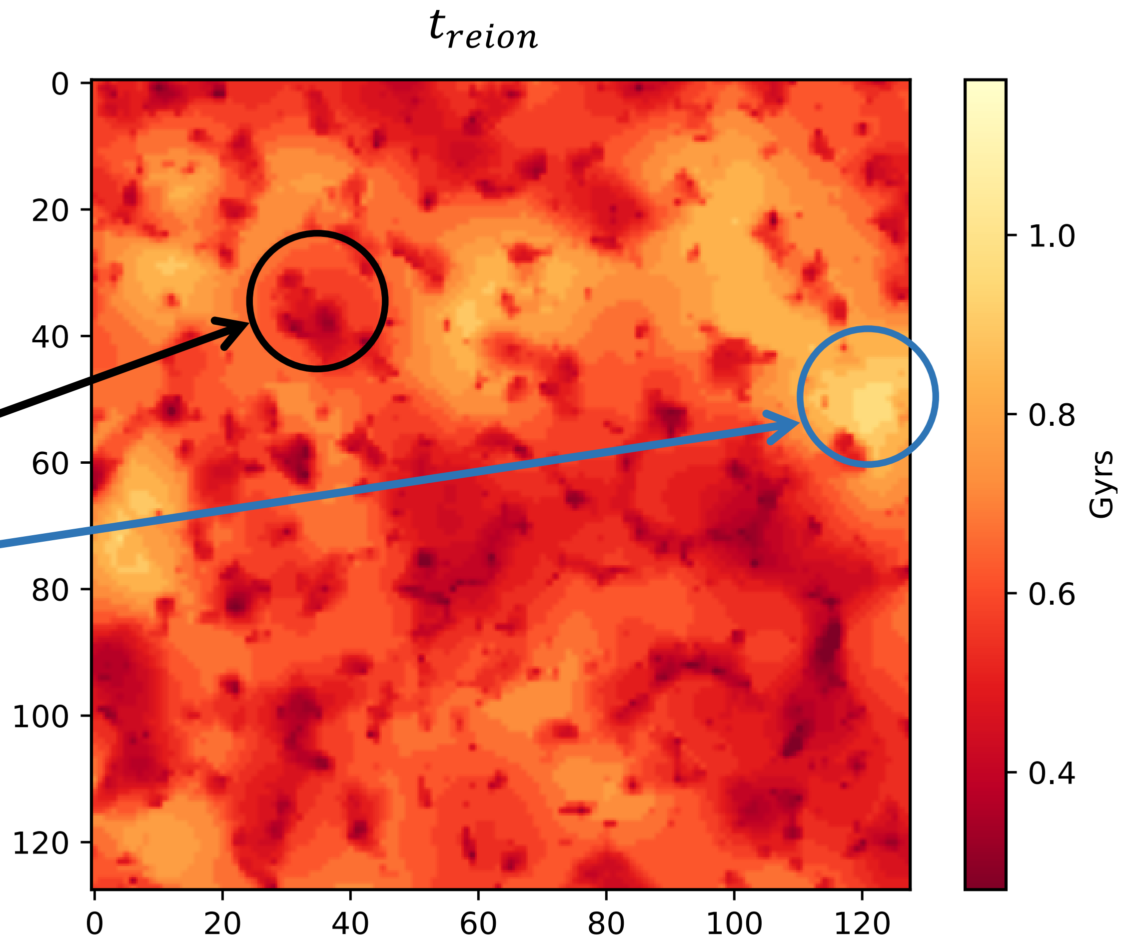
e.g. First/last sources related to minima/maxima of t_{reion} (likely the densest/emptiest regions)

e.g. Isocontour length
=
Typical size of bubbles

Thélie+2022,2023

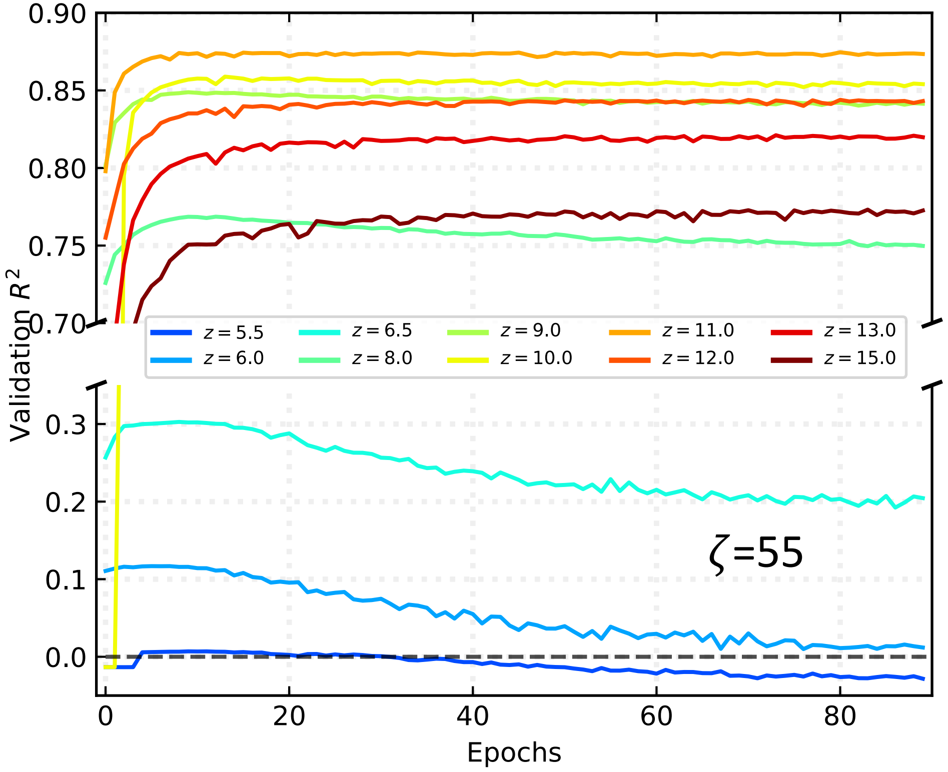
Minima

Maxima



It depicts the history of reionisation of a given region in the sky

MSE – Coefficient of Determination R^2



The determination coefficient is defined as :

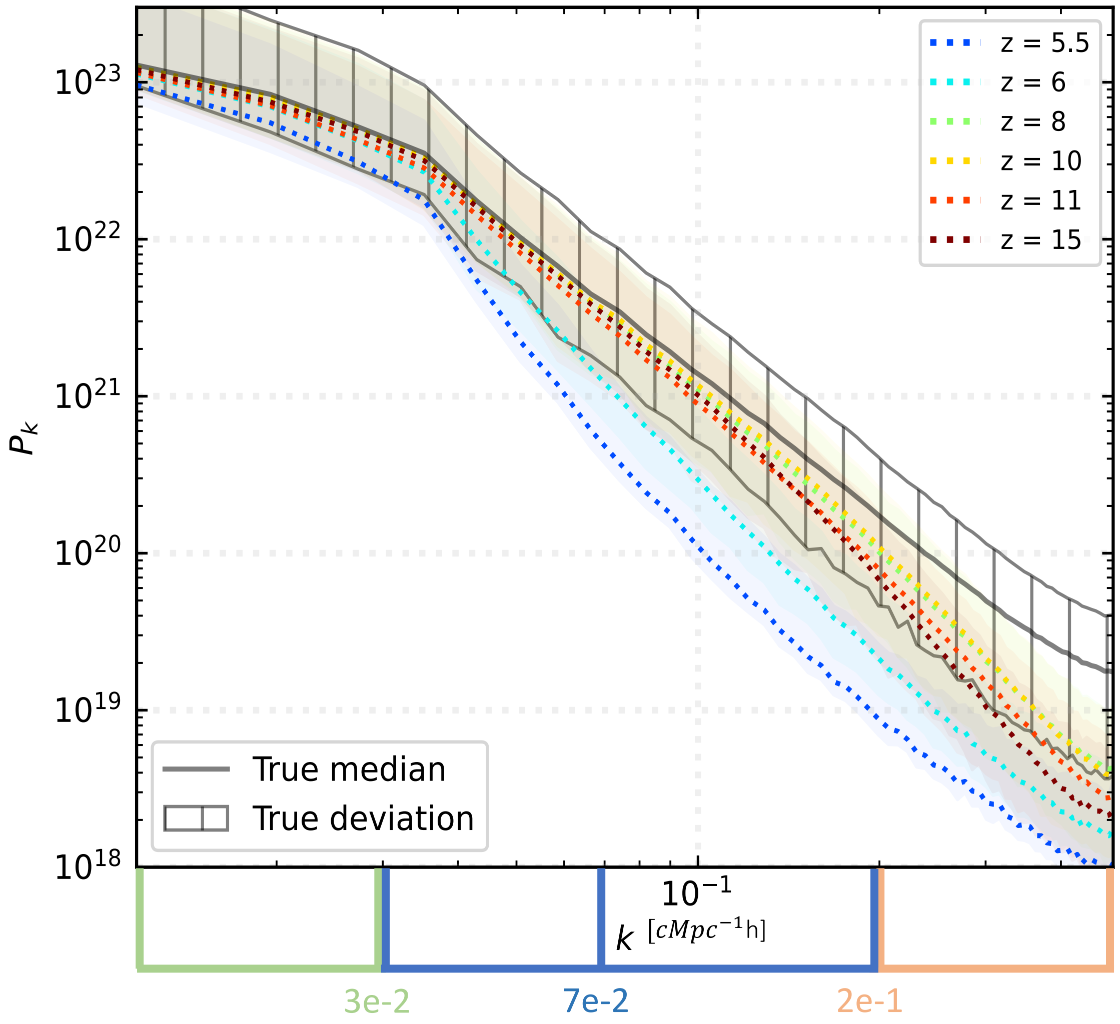
$$R^2 = 1 - \frac{\Sigma(Pred - True)^2}{\Sigma(True - < True >)^2}$$

This coefficient tends to 1 for a perfect correlation between the prediction and the ground truth

<u>R^2 coefficient computed on the validation set:</u>

- Bad performances for low redshift: z < 6.5 ⬄ signal fraction < 40%
- Optimal performance for z = 11
- For z > 11, performances get worse with increasing redshift

2D Power Spectrum of reconstructed t_{reion}



Large Scales k < 3e-2 $cMpc^{-1}$h:
Predictors for all z_{obs} succeed in recovering the larger scales:
more than 95% for z = 8

Intermediate Scales:
Impossibilities to recover the power at this range for lowest z_{obs} :
28% for z = 6 at k = 7e-2 $cMpc^{-1}$h
for Larger z_{obs} the power remaining tends to align with the truth:
85% for z = 8 at k = 7e-2 $cMpc^{-1}$h

Small Scales k > 2e-1 $cMpc^{-1}$h:
Huge difficulties to recover the power for all z_{obs}:
57% for z = 8
12% for z = 6

Fraction of Neutral Hydrogen (Q_{HI})

Reconstruction for observation at **redshift<8** fails at early and late times

z>8 predictors give a relevant history of reionisation.

Cross markers:

computed from **T21 maps** as the signal fraction, a given marker depicts the fraction of HI Q_{HI} at a given redshift using observations at this same redshift

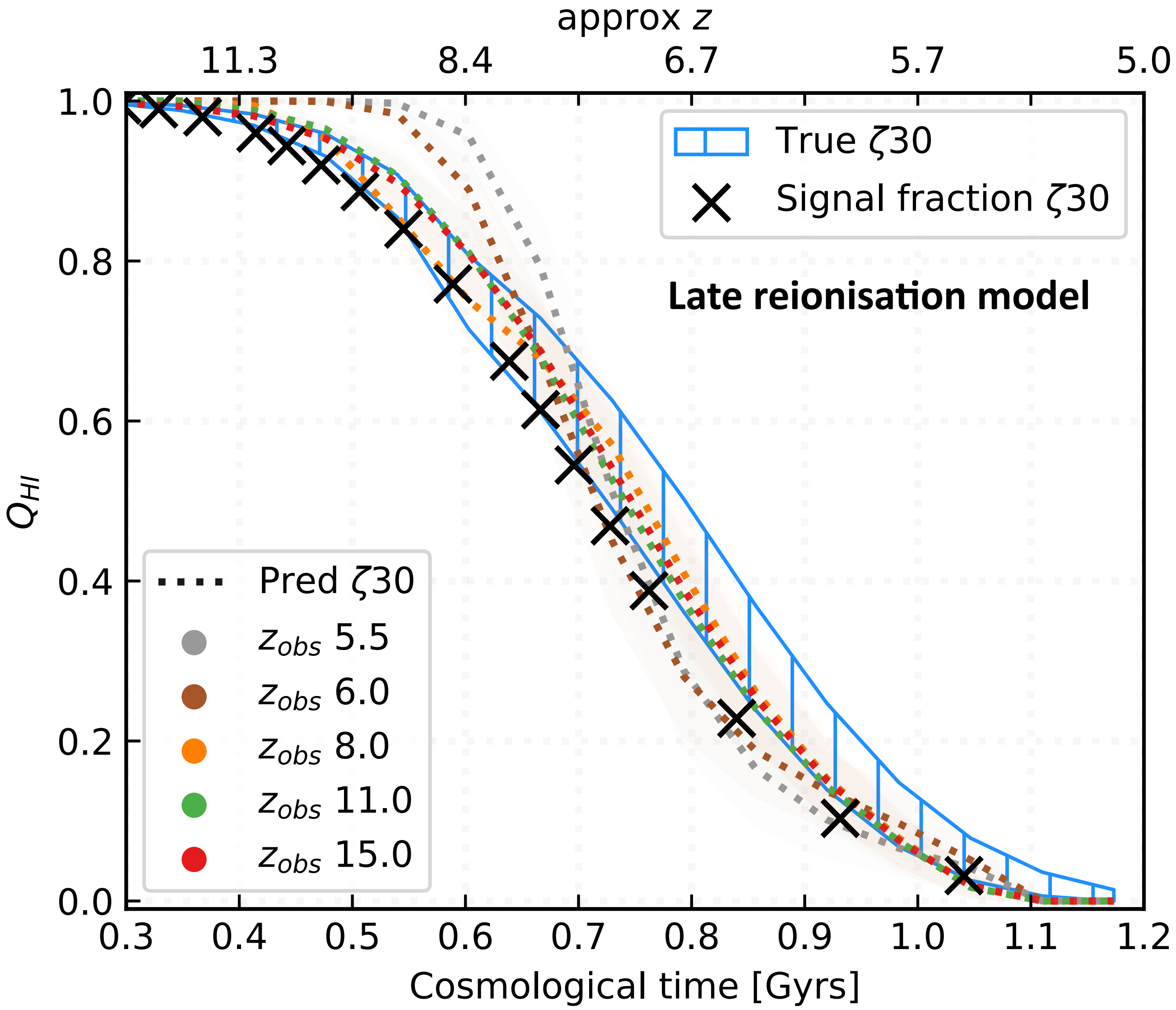
→ **This information is contained in the line of sight**

Dashed area and dotted lines:

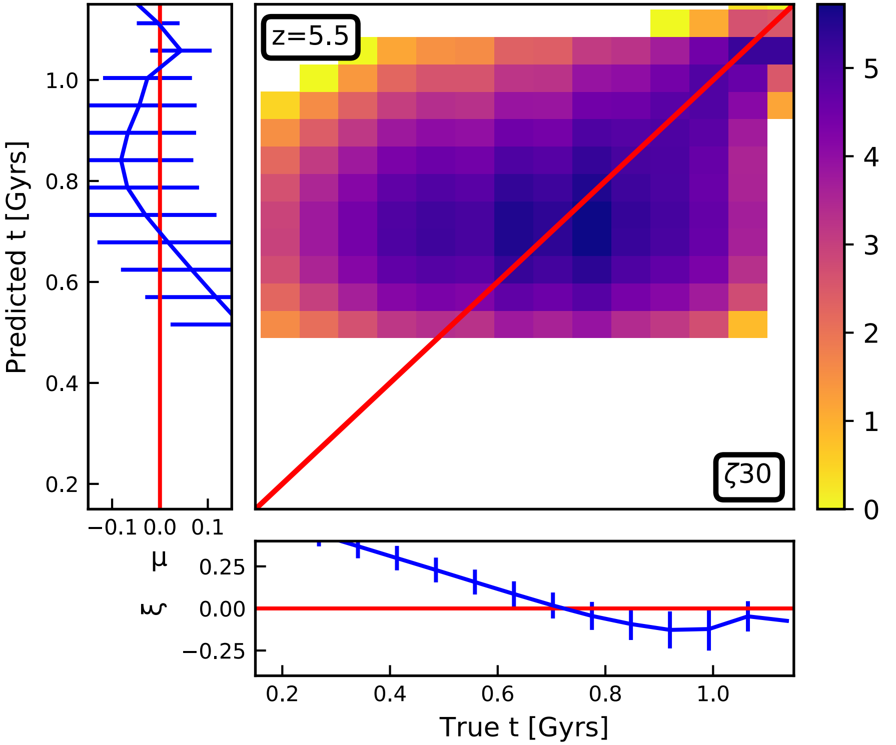
Computed from **time of reionisation maps** as the cumulative PDF

→ **This information is contained in the sky**

Line of sight and plane of the sky are in agreement



True Versus Prediction



Predictor z = 5.5

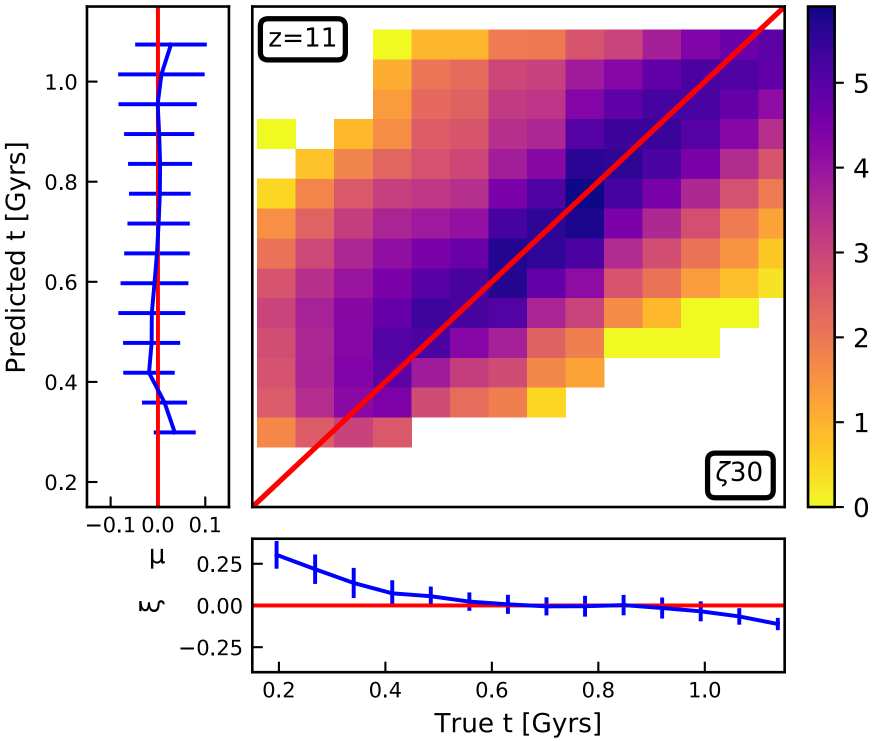
Spread distribution ⇔ predictions doesn't match the truth

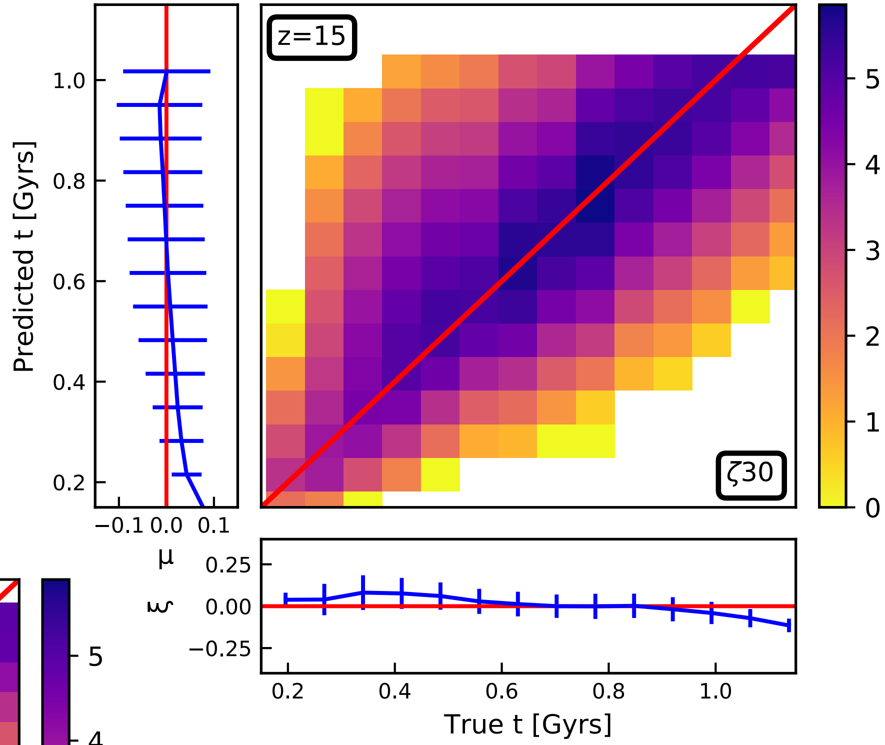
Gets information about last regions to reionise and their value

Predictor z = 11

Major number of predicted pixels fitting with the truth

Missing first sources and last regions to reionise





Predictor z = 15

Major number of predicted pixels fitting with the truth

More information for the first sources than for z = 11

Missing last regions to reionise

WDM Models
21 cm Maps

In WDM models: There is fewer structures than for CDM depending on the mass of the WDM particle

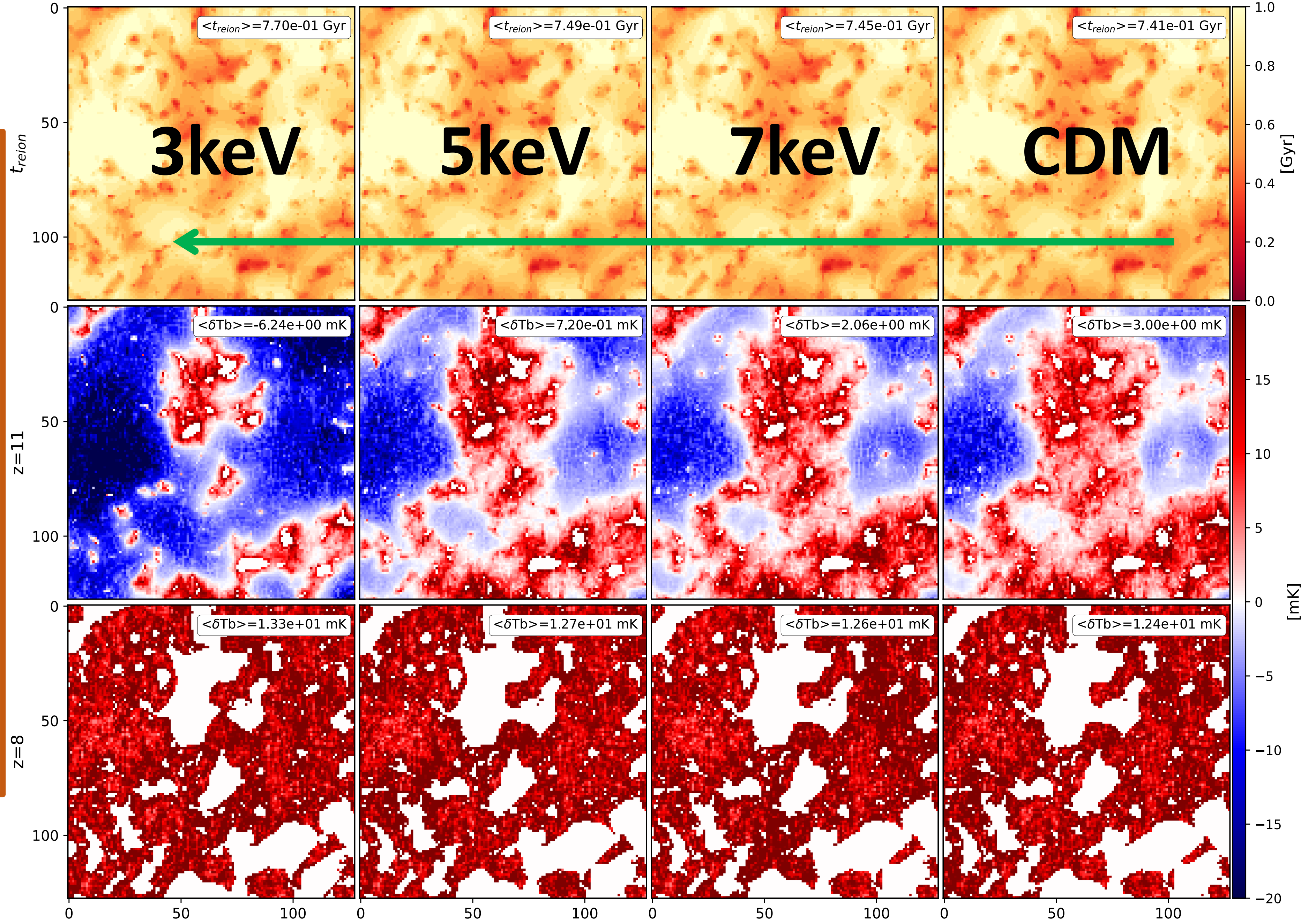
What happens for the 21 cm signal?

21 cm map at z = 11

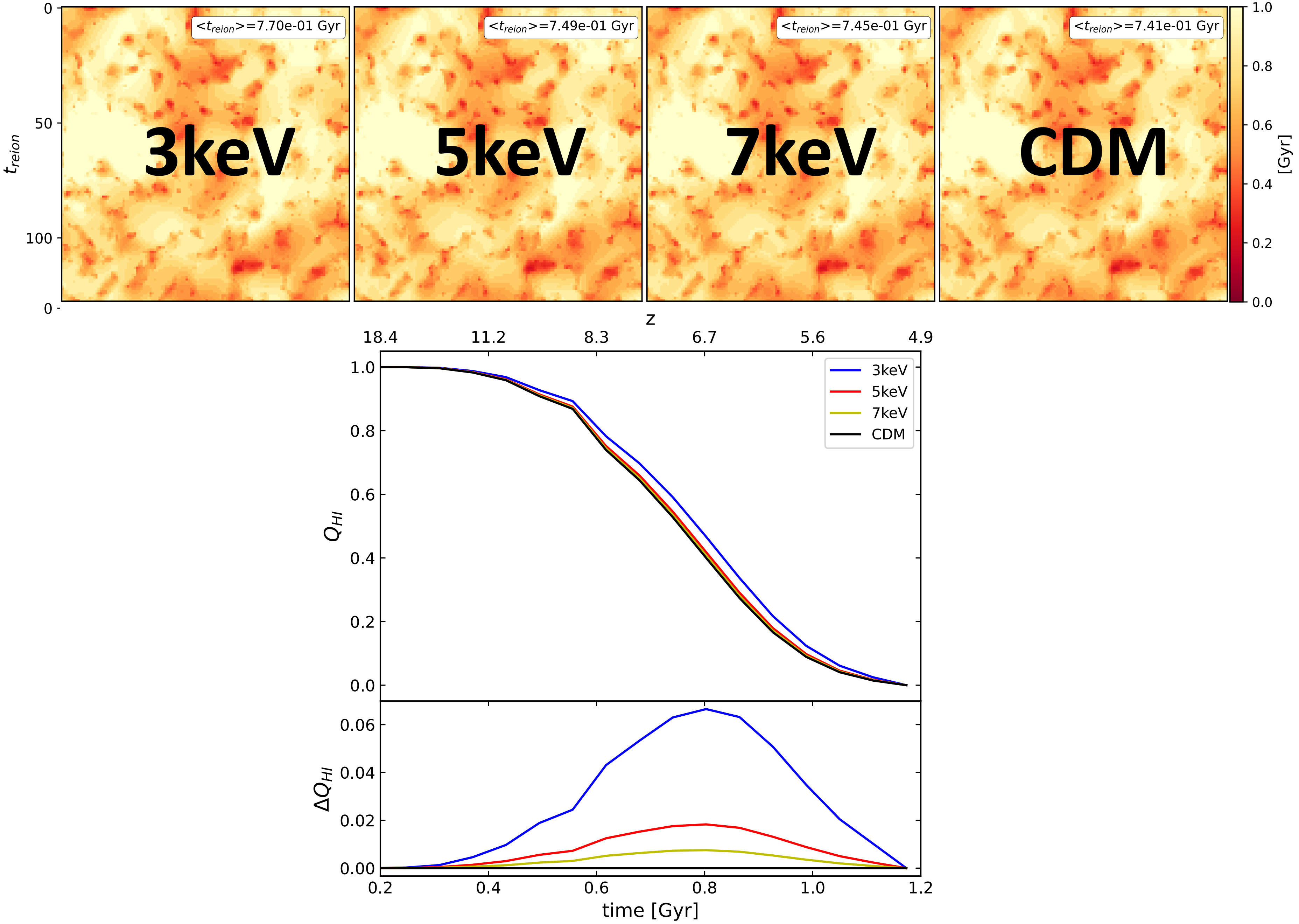
As the **mass** of WDM particle **decreases**:

-> Missing regions

-> In general the distribution of temperature is different because the reionisation timing is different



WDM Models
Reionisation Time Field



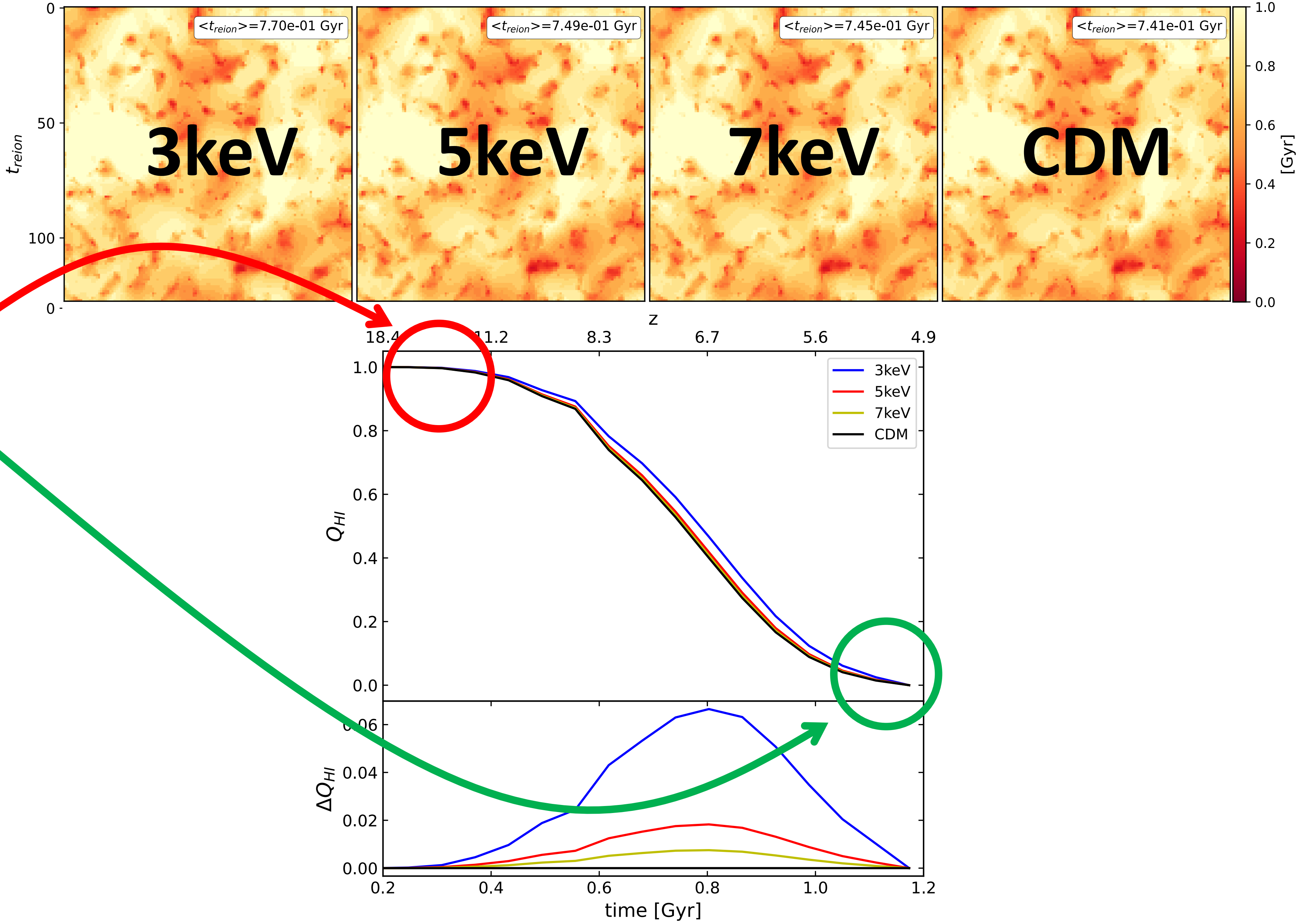
Similar reionisation scenario

Same parameters as previously used for CDM scenario (ζ30) except **for 3keV: ζ = 32**

Reionisation **begins** and **ends** around the same times **z = 15 and 5** respectively

WDM models have a **delayed mean reionisation time**

WDM Models
Reionisation Time Field



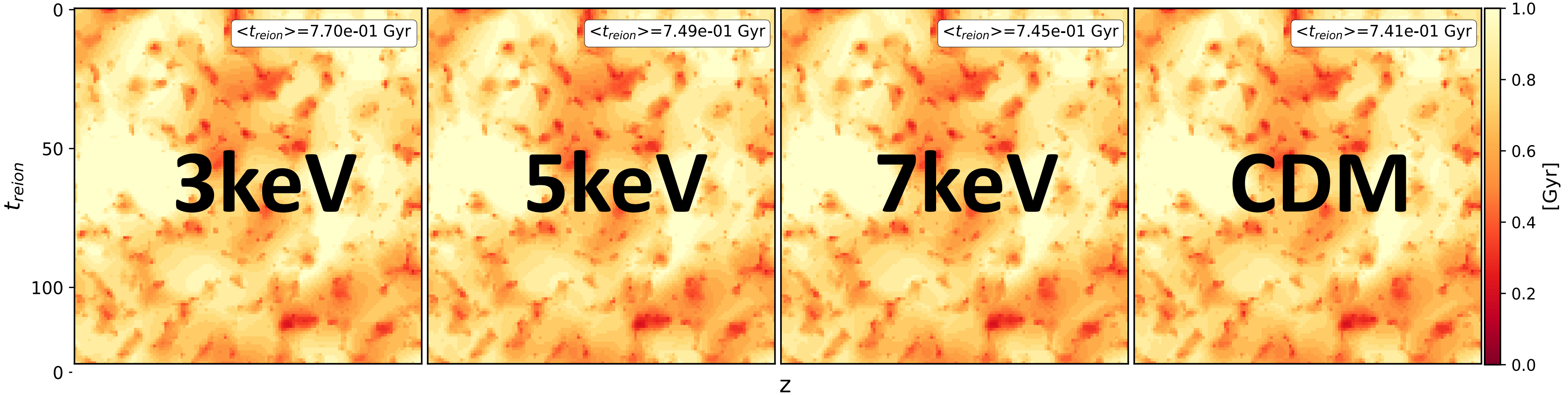
Similar reionisation scenario

Same parameters as previously used for CDM scenario (ζ**30**) except **for 3keV: ζ = 32**

Reionisation **begins and ends** around the same times **z = 15 and 5** respectively

WDM models have a **delayed mean reionisation time**

WDM Models
Reionisation Time Field

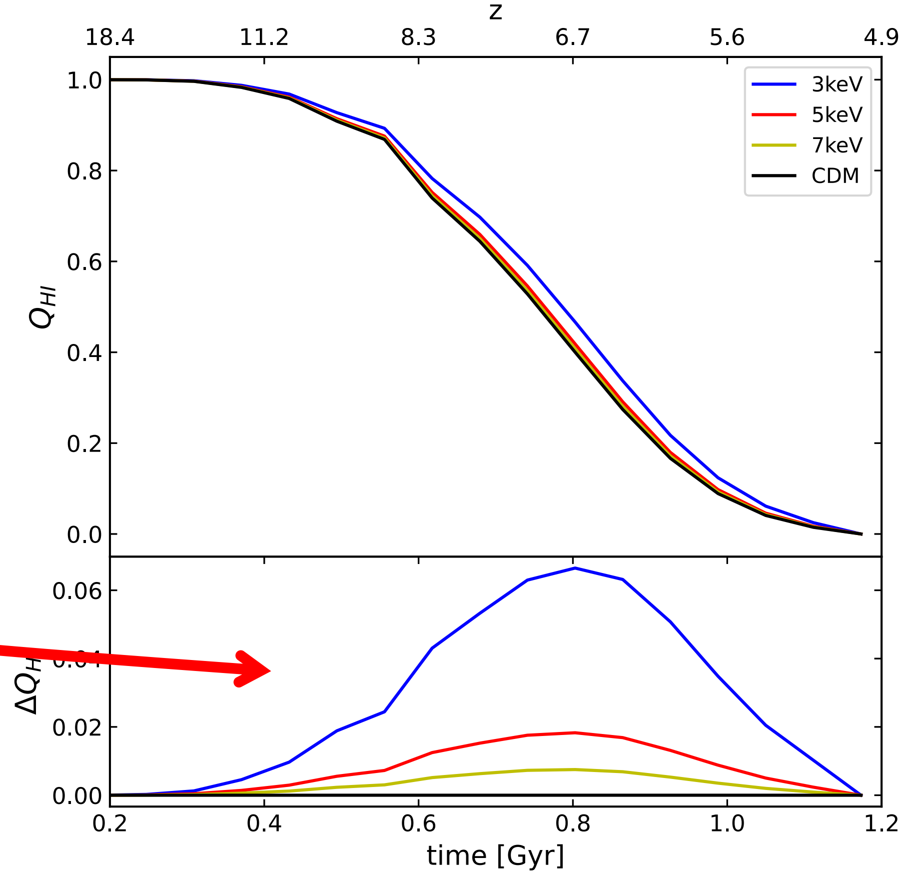


Similar reionisation scenario

Same parameters as previously used for CDM scenario (ζ**30**) except **for 3keV: ζ = 32**

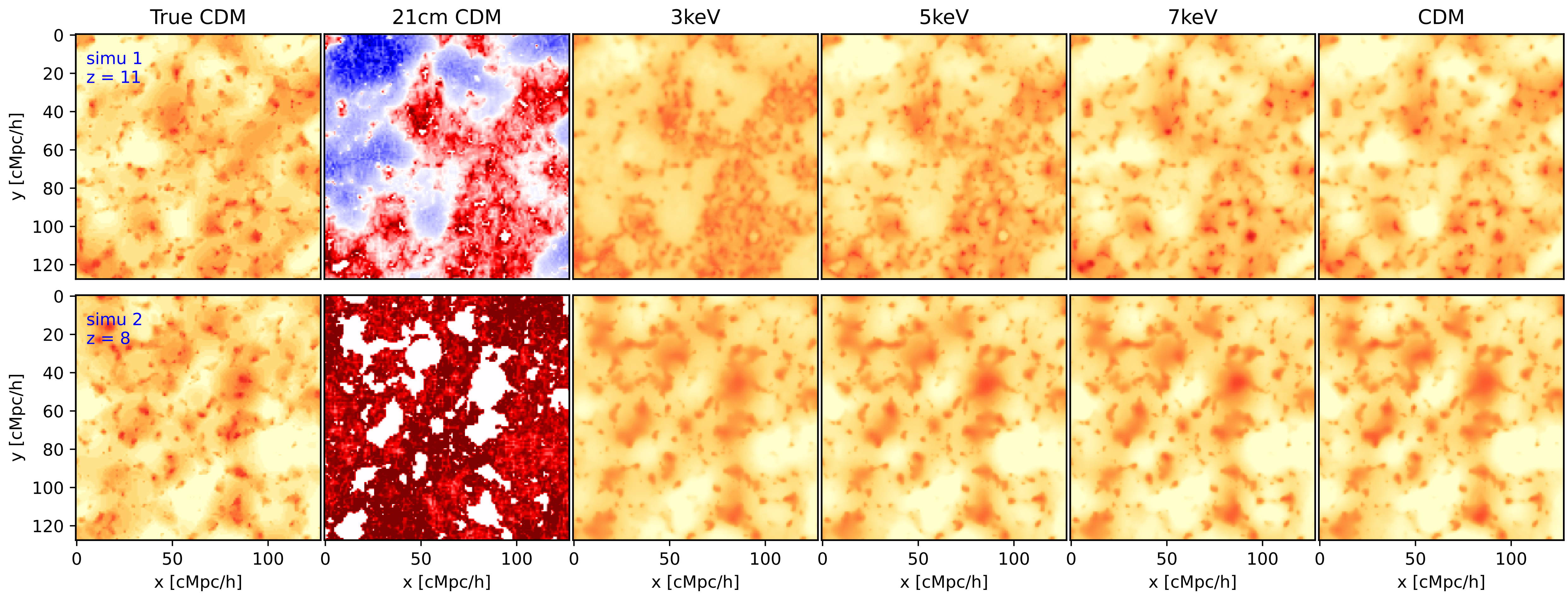
Reionisation **begins and ends** around the same times **z = 15 and 5** respectively

WDM models have a **slight delayed mean reionisation time**

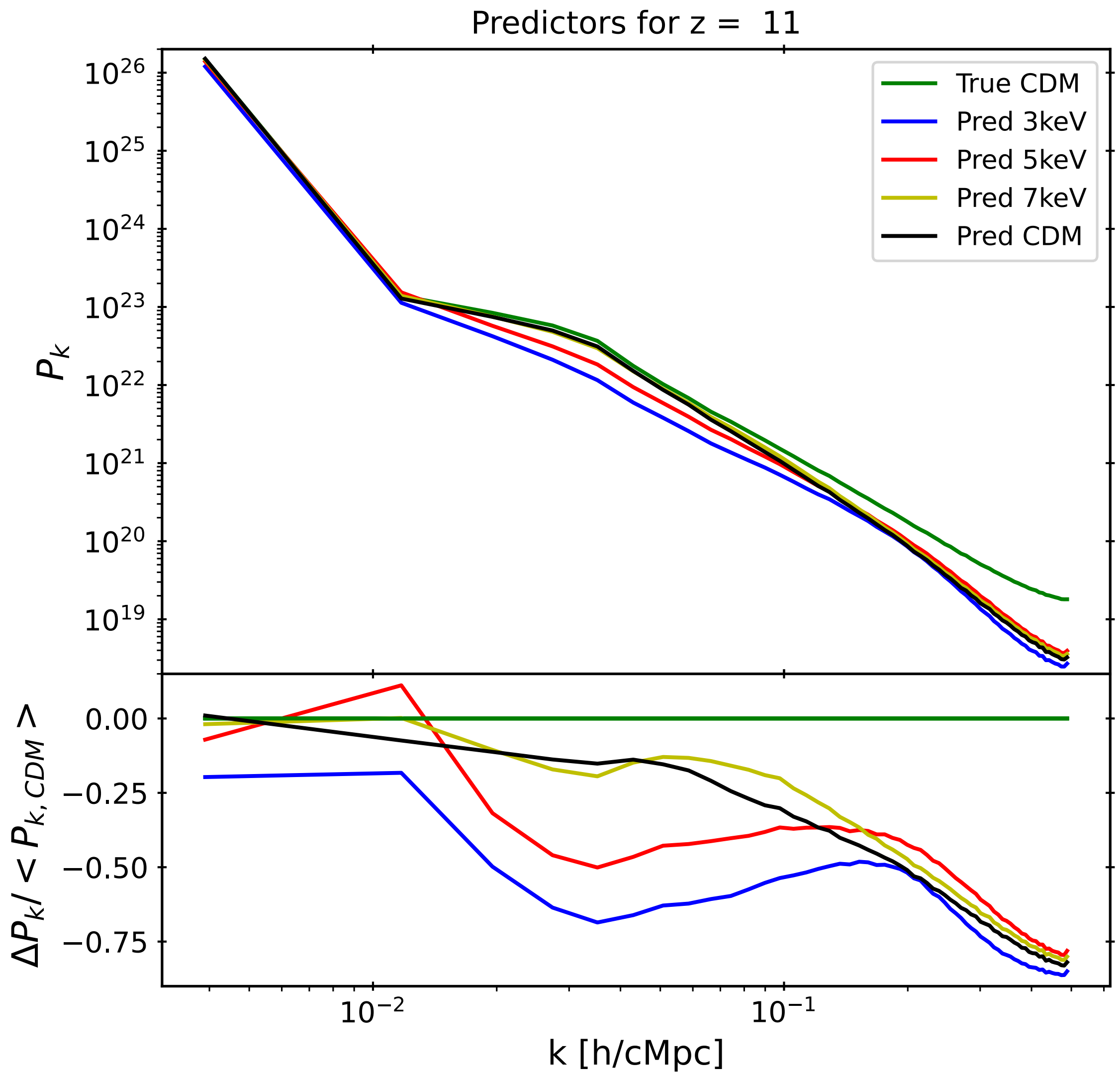


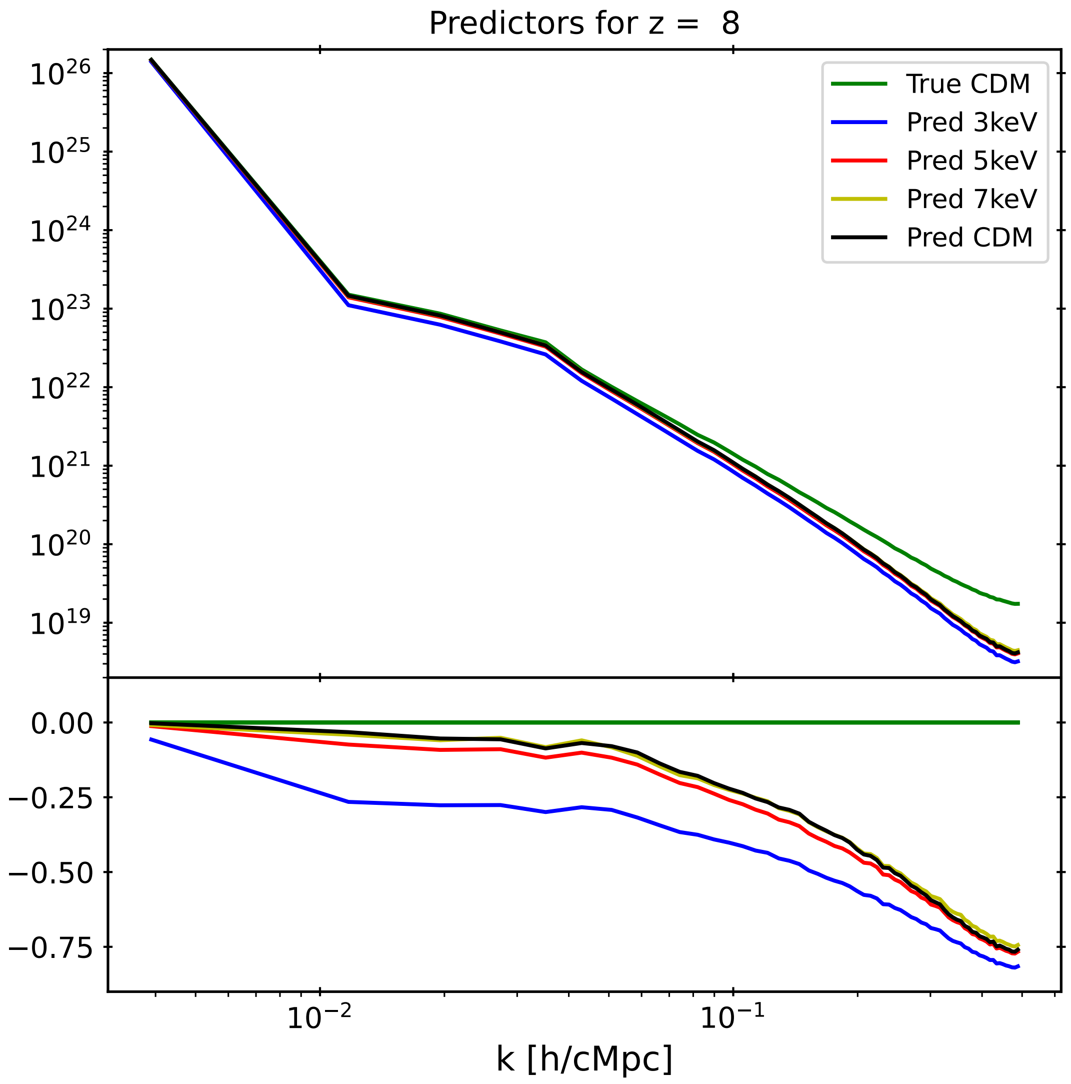
Predictions: CDM vs WDM Predictors

In practice:
- Predictions for z=11 and z=8 **should be statistically similar** → 7keV and CDM look similar
- Predictions for all models should give the **same result** → Not the case for z = 11



Power Spectrum: Comparison to true CDM





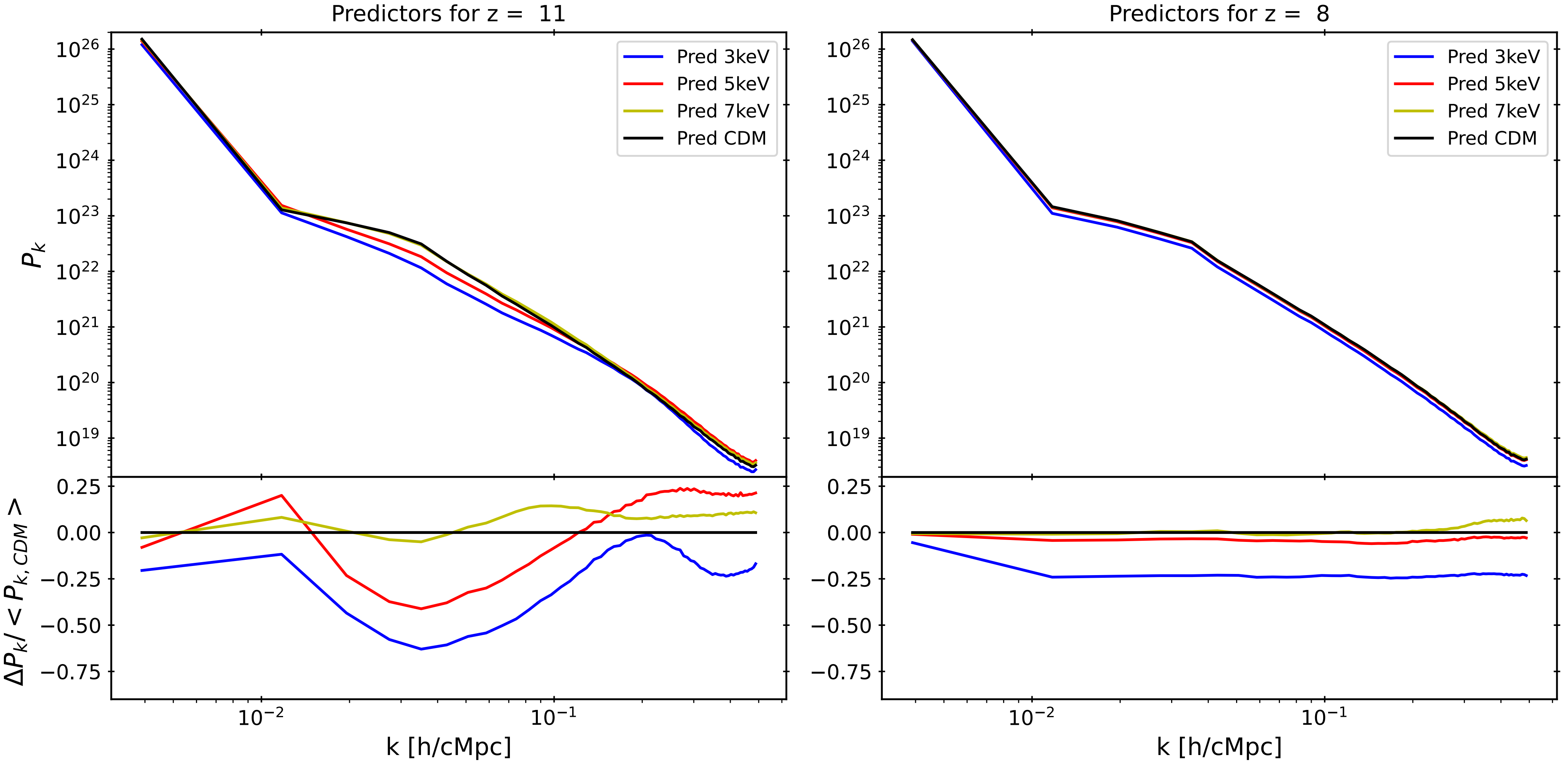
Power Spectrum: WDM vs CDM

We compare **predictions of each model** with predictions of **CDM -> CDM is the reference**, we should get **similar results**.

3keV is way **below the CDM** (>20%) for both redshift: This model can already be **excluded** from this preliminary study.

5keV behave weirdly for z =11: it is a first hint but it is not enough to exclude it

7keV is close to CDM: cannot conclude yet

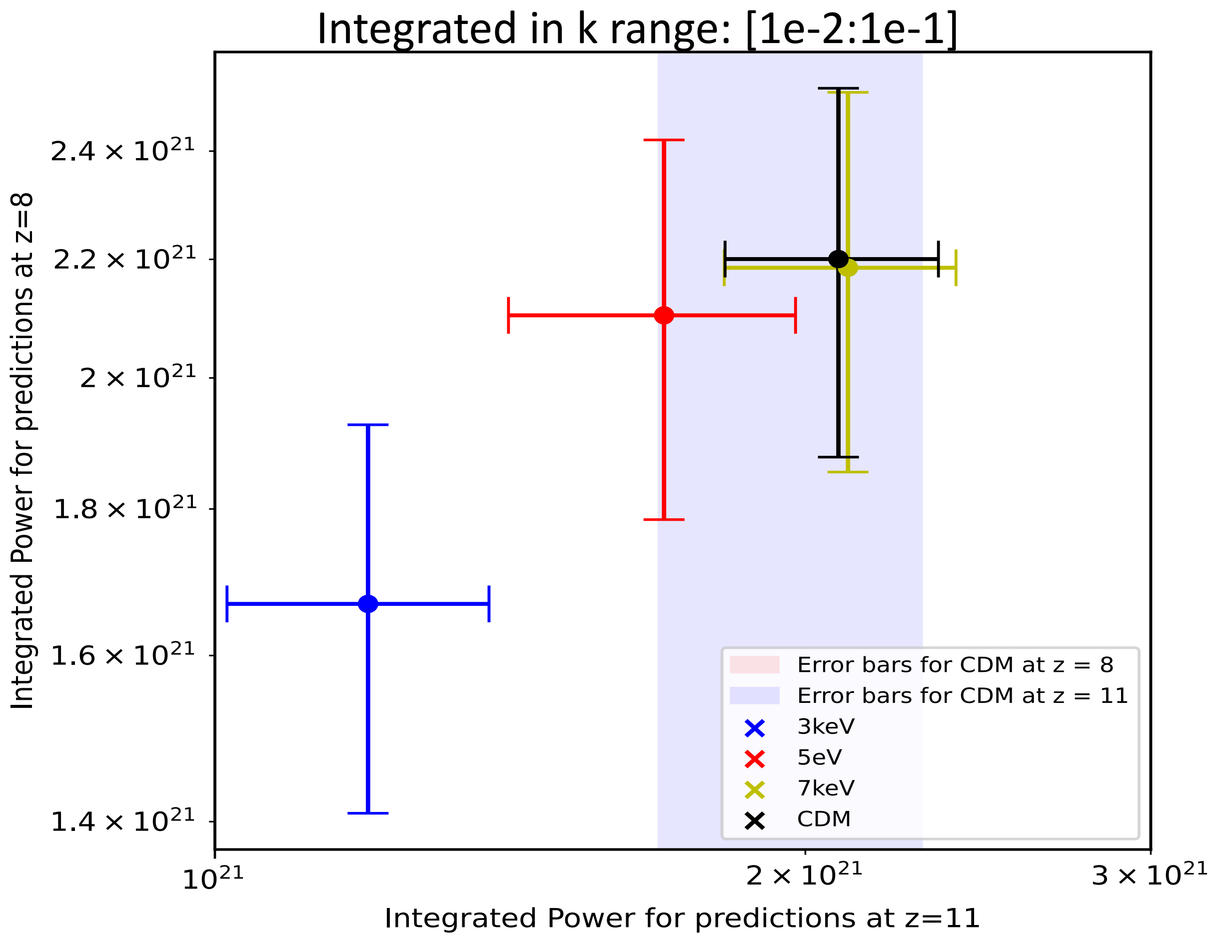


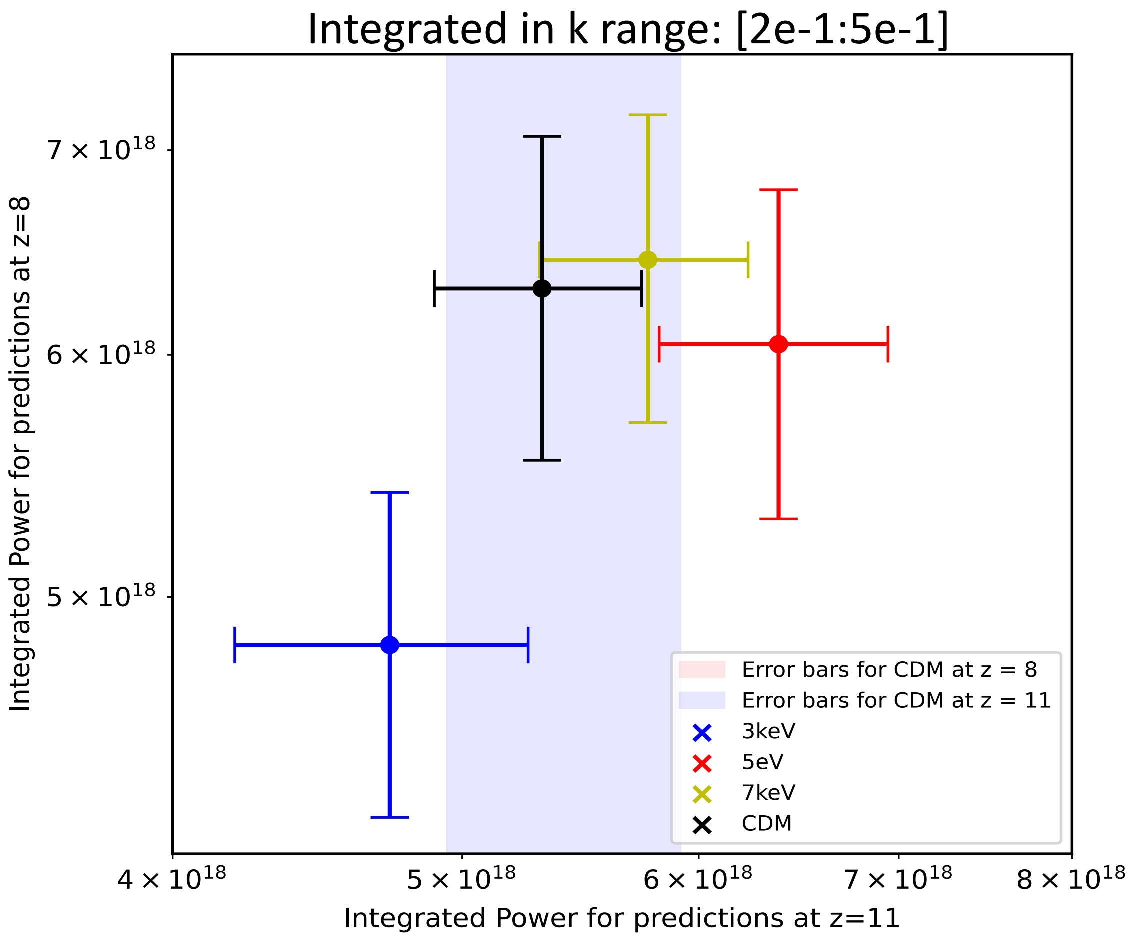
Integrated Power

From CDM predictions: We can obtain the mean and std of the power spectrum and **integrate it for a given k range**:

–> We obtain our **reference** where the predictions **should be located**. (**red and blue bands**)

We perform the same transformation for the predictions of each model

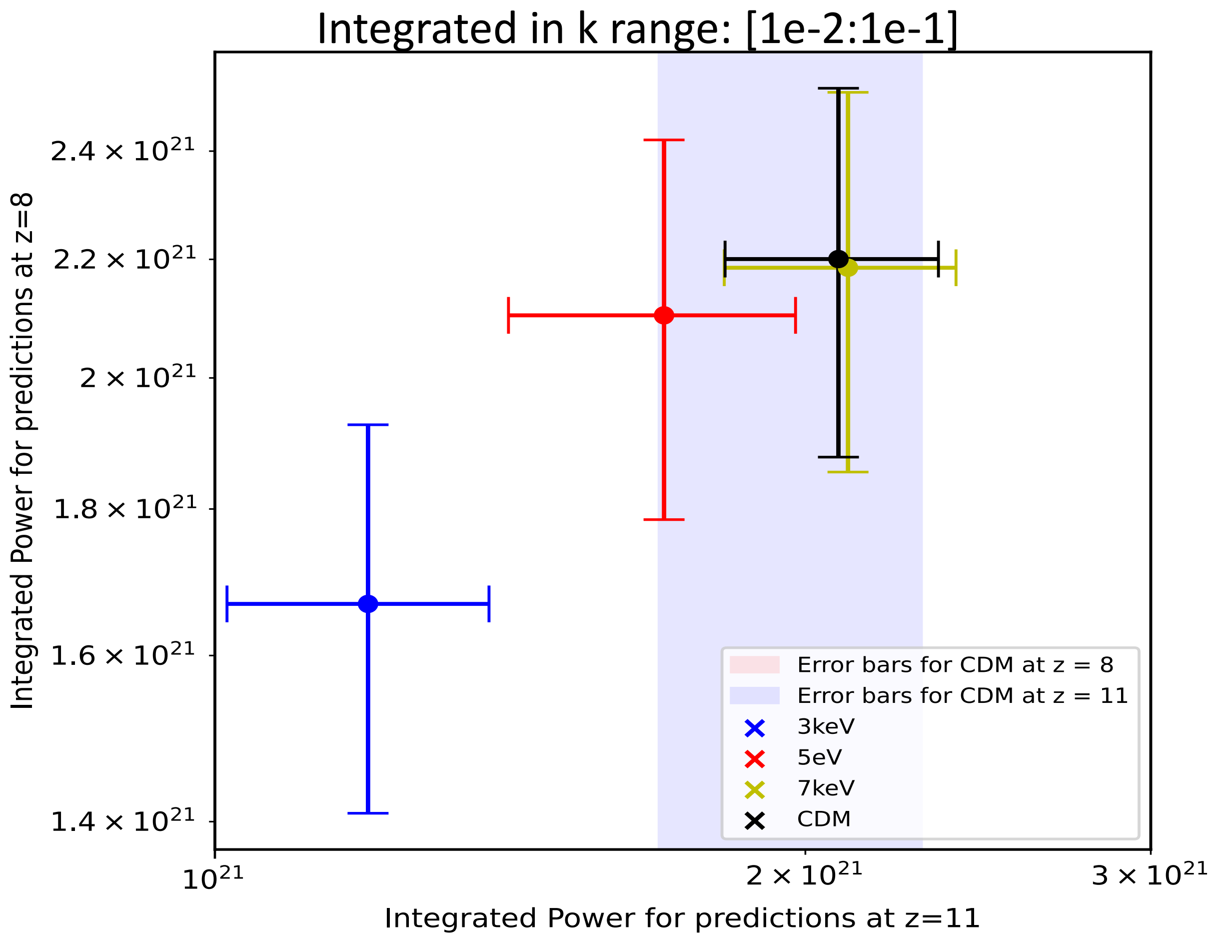


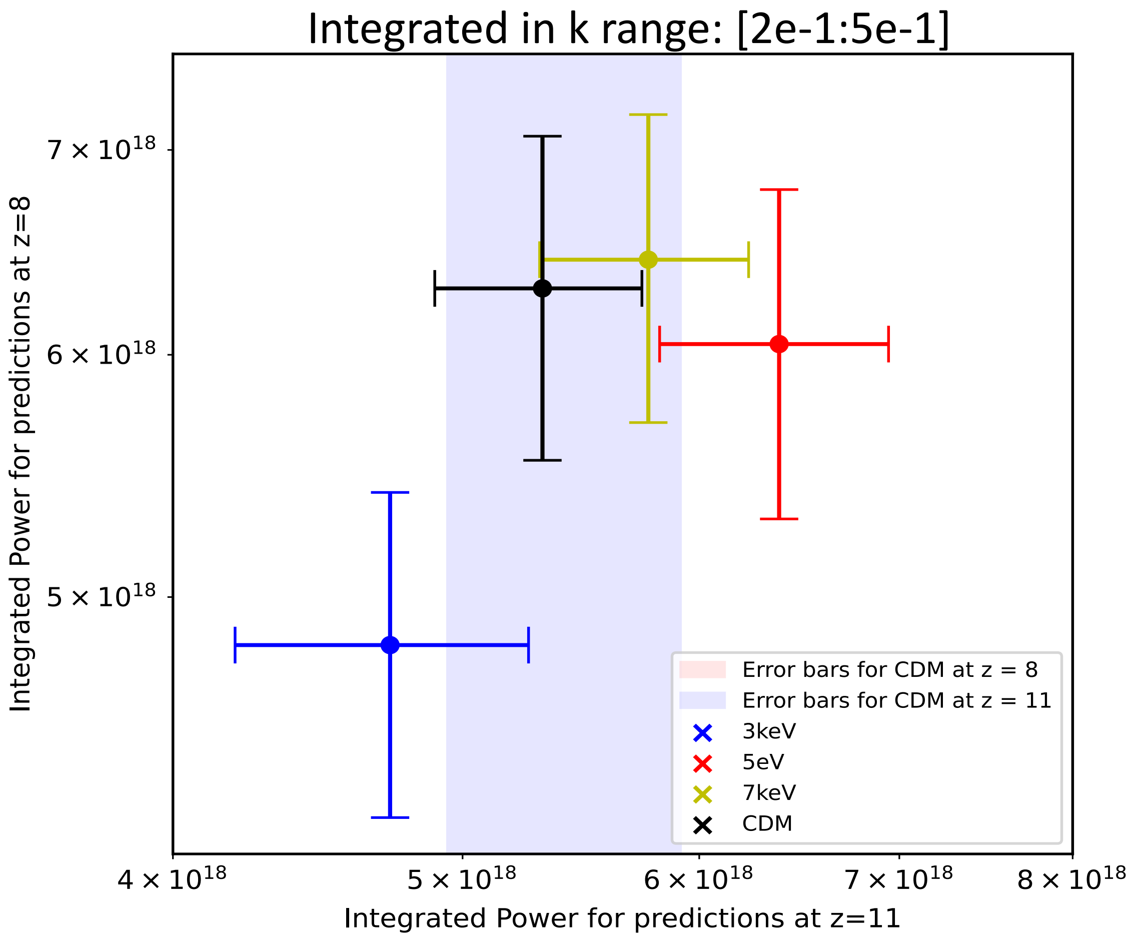


Integrated Power

CDM matches with the reference (Fortunately, the reference is the CDM !)
7keV matches as well with the reference: Cannot conclude anything
5keV matches when z=8 but differs for z=11: Need to confirm with another metric
3keV is off limits: This model is **excluded**

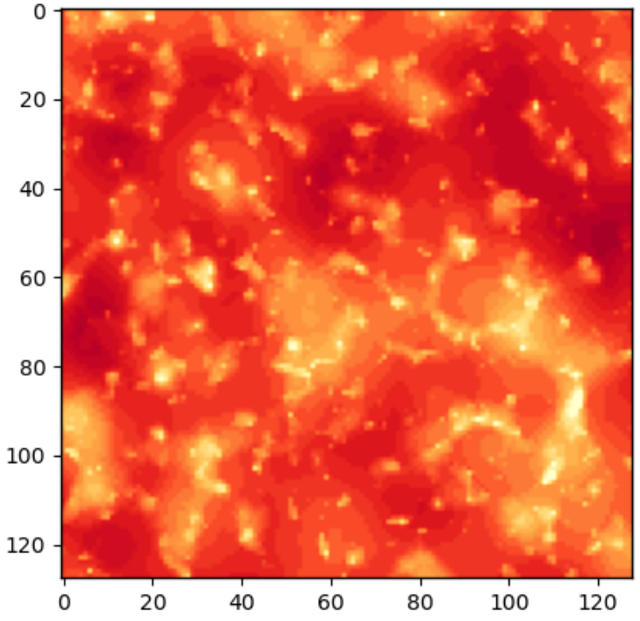
Issue with this metric:
What is the **meaning** of an **integrated power**?
Results are **dependent** of the **k range chosen**



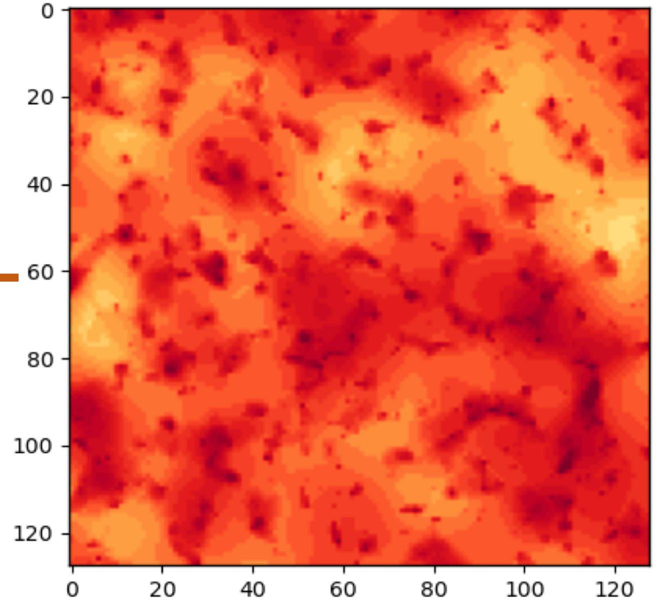


21cmFast simulation code: z_{reion} Vs t_{reion}

z_{reion}

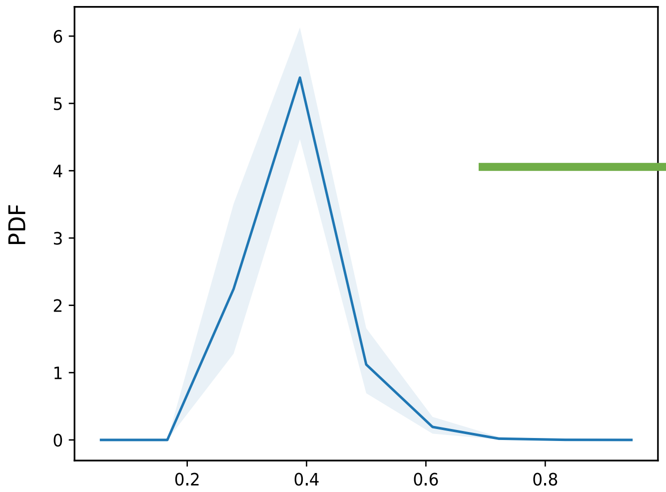


t_{reion}

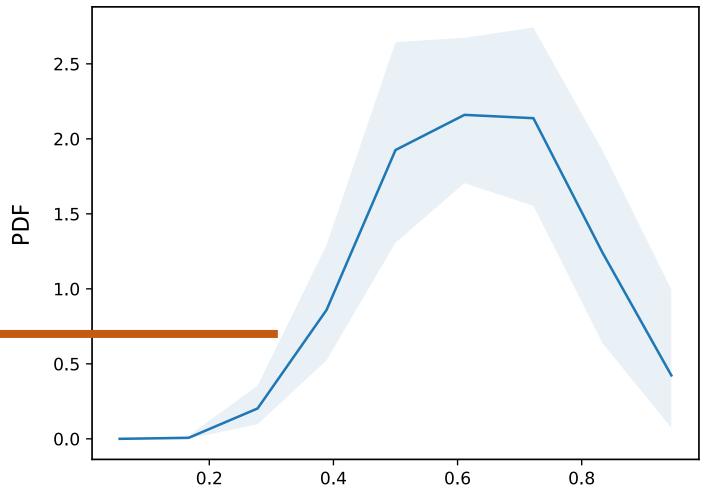


More contrast

Less contrast
Smoother

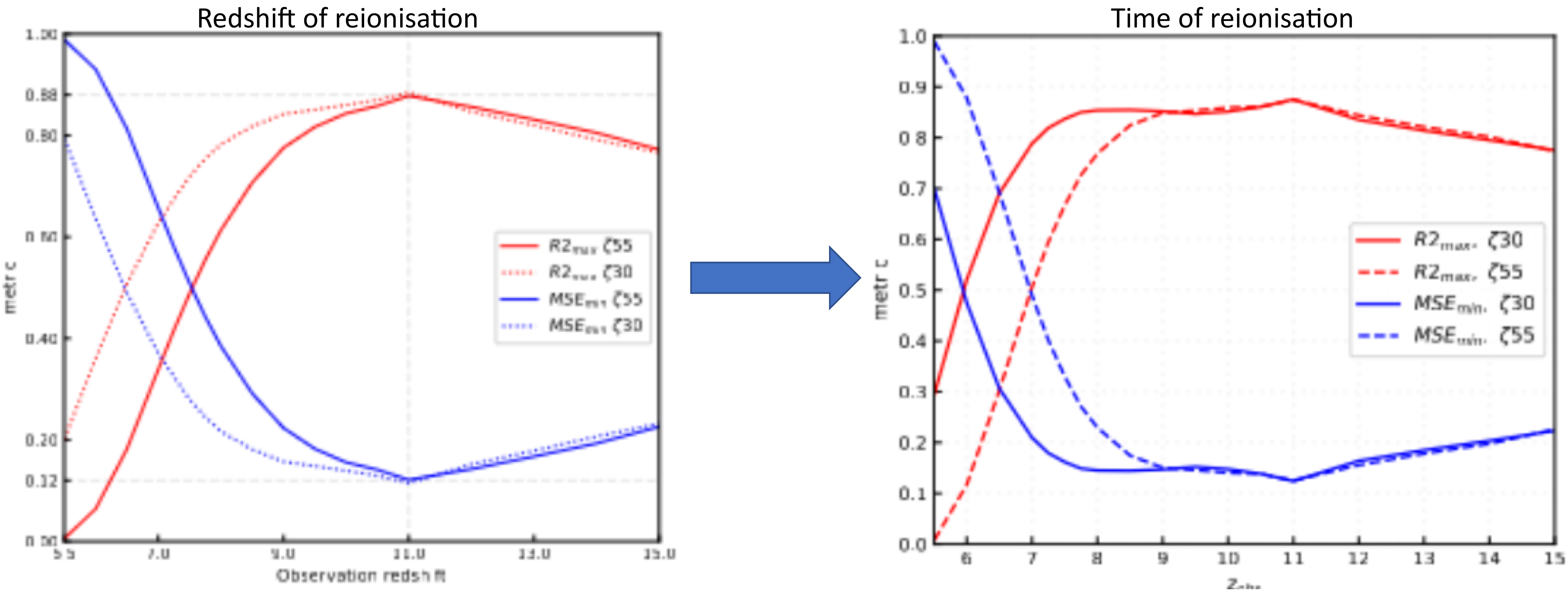


Tighter

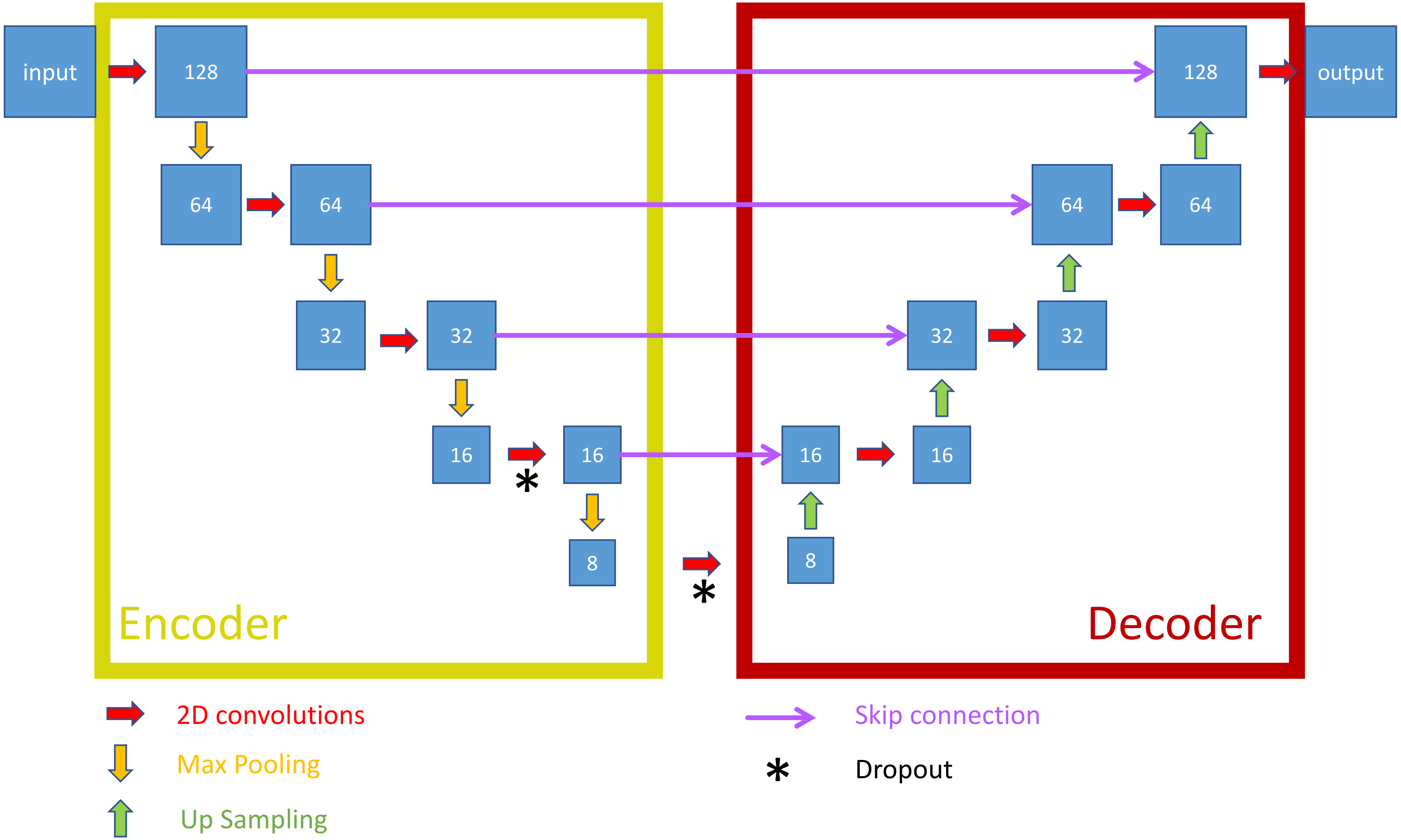


Larger distribution
More symmetrical
More gaussian

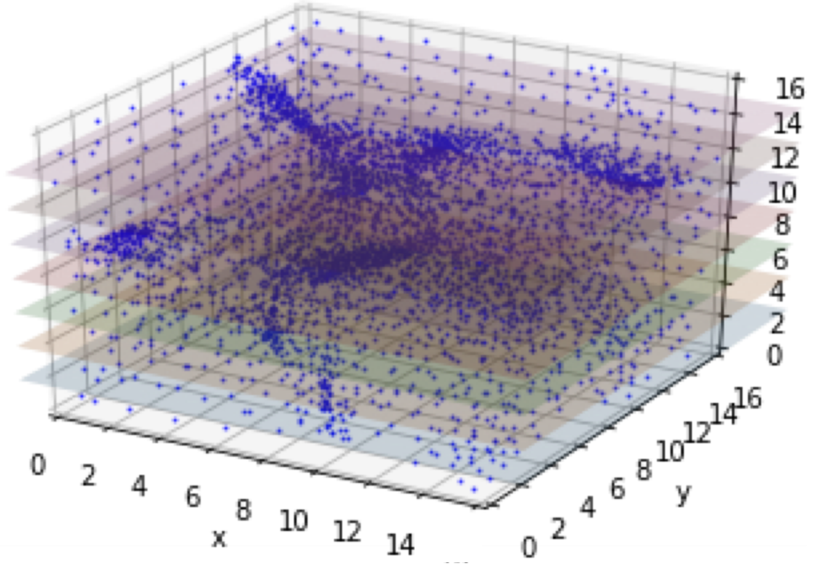
21cmFast simulation code: z_{reion} Vs t_{reion}



Convolutional Neural Network: U-net Architecture



21cmFast simulation code: image extraction process



Example of the process to extract Images from a box of size 16^3 using a spacing of 2 pixels

For a single simulation (= 1 box of volume 256^3):

- Divide it into 8 cubes of size 128^3

For each of cube:

- Extract 32 evenly spaced (4 pixels) slices (128^2) in each direction (x,y,z)

Resulting in 96 images for each cube of size 128^3
Leading to 768 images for 8 cubes (= 1 simulation)

To account for different redshifts used in the training of the CNN, this number is multiplied by 18

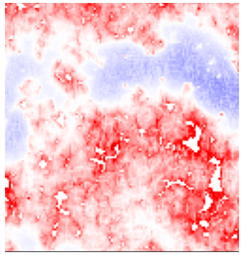
There are 2 cosmological models, further increasing the total number of images by a factor 2

The dataset used in this work consists of **50 simulations**, resulting in a total of **38 400 images / field (x2) / redshift (x18) / reionisation model (x2)**

Note: for t_{reion} we don't multiply by 18

Convolutional Neural Network: **Training Phase**

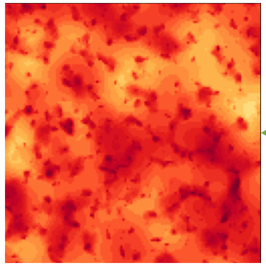
Input



31 500 maps
At a given
redshift

CNN

Output



31 500 maps

= 1 predictor

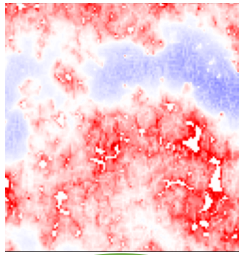
Predictions of the algorithm are not yet in their final form since they will continue to evolve until the learning phase is done

The update of the weight happens each time a batch of input passes through the CNN by computing the error (MSE) done comparing its prediction with ground truth

1 predictor / redshift and / model

Convolutional Neural Network: **Validation Phase**

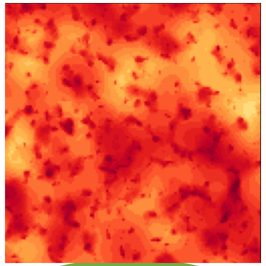
Input



3 500 maps
At a given redshift

CNN
In training process

Output



3 500 maps

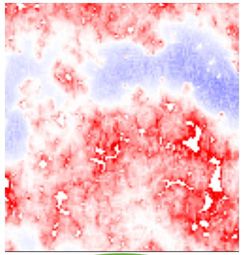
These maps are distinct from maps in the training set

During the Validation phase that takes place after each epoch, the algorithm compute the MSE and the R^2 coefficient on images that the predictor has never seen

Nevertheless, the weights are not updated at this step

Convolutional Neural Network: **Testing Phase**

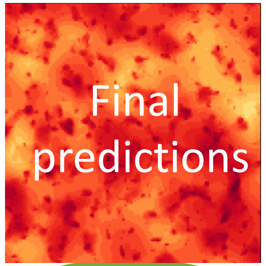
Input



3 400 maps
At a given
redshift

Predictor
At a given
redshift

Output



3 400 maps

These maps are distinct from maps in the training and validation set

The testing phase takes place when the predictors are set. It is now used to predict the time of reionisation from the temperature brightness.

We now can use the predicted fields to measure parameters and compare it via several metrics to the ground truth given by the simulation