

A semi-analytic model comparison: testing cooling models against hydrodynamical simulations

Pierluigi Monaco,^{1,2★} Andrew J. Benson,³ Gabriella De Lucia,² Fabio Fontanot,^{4,5} Stefano Borgani^{1,2,6} and Michael Boylan-Kolchin^{7,8}

¹Dipartimento di Fisica – Sezione di Astronomia, Università di Trieste, via Tiepolo 11, I-34131 Trieste, Italy

²INAF, Osservatorio Astronomico di Trieste, Via Tiepolo 11, I-34131 Trieste, Italy

³The Observatories of the Carnegie Institution for Science, 813 Santa Barbara St, Pasadena, CA 91101, USA

⁴Heidelberger Institut für Theoretische Studien (HITS), Schloss-Wolfsbrunnengasse 35, D-69118 Heidelberg, Germany

⁵Institut für Theoretische Physik, Philosophenweg 16, D-69120, Heidelberg, Germany

⁶INFN-National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy

⁷Center for Cosmology, Department of Physics and Astronomy, 4129 Reines Hall, University of California, Irvine, CA 92697, USA

⁸Department of Astronomy, University of Maryland, College Park, MD 20742, USA

Accepted 2014 March 31. Received 2014 February 24; in original form 2013 August 28

ABSTRACT

We compare predictions of cooled masses and cooling rates from three stripped-down semi-analytic models (SAMs) of galaxy formation with the results of N -body+Smoothed Particle Hydrodynamics (SPH) simulations with gas particle mass of $3.9 \times 10^6 h^{-1} M_\odot$, where radiative cooling of a gas of primordial composition is implemented. We also run a simulation where cooling is switched on at redshift ~ 2 , in order to test cooling models in a regime in which their approximations are expected to be valid. We confirm that cooling models implemented in SAMs are able to predict the amount of cooled mass at $z = 0$ to within ~ 20 per cent. However, some relevant discrepancies are found. (i) When the contribution from poorly resolved haloes is subtracted out, SAMs tend to underpredict by ~ 30 per cent the mass that cools in the infall-dominated regime. (ii) At large halo masses, SAMs tend to overpredict cooling rates, though the numerical result may be affected by the use of a standard version of SPH. (iii) As found in our previous work, cooling rates are found to be significantly affected by model details: simulations disfavour models with large cores and with quenching of cooling at major mergers. (iv) When cooling is switched on at $z \sim 2$, cold gas accumulates very quickly in the simulated haloes. This accumulation is reproduced by SAMs with varying degrees of accuracy.

Key words: methods: numerical – galaxies: evolution – galaxies: formation.

1 INTRODUCTION

The formation of galaxies within the Λ cold dark matter (Λ CDM) cosmological model involves a large number of physical processes, many of which are still poorly understood. The hierarchical build-up of dark matter (DM) haloes, resulting from the non-linear evolution of primordial perturbations under their own gravity, provides the backbone of the whole process of galaxy formation. Thanks to advances in N -body techniques and to the very accurate constraints available on cosmological parameters (e.g. Planck Collaboration 2013), the evolution and properties of DM haloes can be computed with very good accuracy (e.g. Reed et al. 2013). Despite baryonic processes are known to affect the build-up of DM haloes to some extent (e.g. Stanek, Rudd, & Evrard 2009; Duffy et al. 2010; Saro

et al. 2010; Cui et al. 2012), the modelling of baryonic physics still provides the major source of uncertainty. A purely collisionless simulation thus remains a good starting point for a galaxy formation model.

Galaxy formation in the cosmological context has been historically addressed with two main tools. Semi-analytic models (SAMs) are applied to the backbone of DM halo merger trees, taken from an N -body simulation or equivalent tools (e.g. Monaco et al. 2013). They use a set of simplified or phenomenological models to describe the various processes that involve baryons. Hydrodynamic simulations consist of numerically solving the equations of motion of DM and gas particles in a realization of a cosmological volume. Two fundamental processes take place on scales that are within reach of presently available simulations. Gravity and hydrodynamic forces are responsible for infalling of gas and heating to the halo virial temperature. Radiative cooling down to $\sim 10^4$ K, which is computed based on the density and temperature of the heated gas, is

★ E-mail: monaco@oats.inaf.it

responsible for the condensation of gas into the inner regions of the halo. There, the gas can fragment into stars and thus form a galaxy, but star formation and all the processes triggered by it (stellar feedback, chemical evolution, galaxy winds, black hole seeds), not to mention the accretion of gas on to black holes, take place at much smaller scales. To properly follow these processes, the range of scales that must be resolved (from sub-pc scale of star formation to cosmological scales) is so vast and the involved physics so complex that their effects need to be treated through simplified, subresolution models.

The modelling of cooling in SAMs is based on the assumption that the gas settles into a hot atmosphere in hydrostatic equilibrium within the potential well of the DM halo. This allows the computation of a cooling time as a function of radius. Whenever the central cooling time is longer than the halo dynamical time ('cooling-dominated regime'; Binney 1977; Rees & Ostriker 1977; White & Frenk 1991), the deposition of cold gas into the central galaxy is assumed to be regulated by cooling; otherwise, the time-scale for gas to condense into the galaxy is assumed to be of the order of the halo dynamical time ('infall-dominated regime'). Using hydrodynamical simulations that included radiative cooling and star formation (but no efficient stellar feedback), Kereš et al. (2005) reported that at $z \gtrsim 2$, or at any redshift for haloes smaller than $\sim 10^{11} M_{\odot}$ (see also Dekel & Birnboim 2006), gas tends not to shock to the virial temperature but to condense directly into the galaxy via a cold flow. As a caveat, the deposition of gas through cold flows is known to depend on the hydrodynamic scheme (Nelson et al. 2013) and to be affected by feedback processes connected to galaxy formation (Benson & Bower 2011; Murante et al. 2012).

To check to what level the two techniques, SAMs and hydrodynamical simulations, give a consistent description of the deposition of cold mass into the 'galaxies', many authors (Benson et al. 2001; Yoshida et al. 2002; Helly et al. 2003; Cattaneo et al. 2007; Viola et al. 2008; Saro et al. 2010; Lu et al. 2011; Hirschmann et al. 2012) performed comparisons of the predictions of SAMs and simulations. All of these papers presented comparisons performed using stripped-down SAMs, where all processes beyond gas cooling were switched off, and simulations, where either no star formation or no (effective) feedback from star formation was present. The first papers (Benson et al. 2001; Yoshida et al. 2002; Helly et al. 2003; Cattaneo et al. 2007) compared one SAM with simulations and reported that SAMs are able to reproduce the gas mass that cools in DM haloes to a level, to cite Benson et al. (2001), '*better than a pessimist might have expected*'. More recent papers focused on some discrepancies between the two kinds of modelling. Viola et al. (2008) simulated cooling in isolated, hydrostatic DM haloes and compared the resulting cooling mass with two models: an implementation from Cole et al. (2000) and the one used in the MOf the Rise of Galaxies MORGANA; And Agn (Monaco, Fontanot, & Taffoni 2007). They found that the former model underestimates the amount of cooled mass when cooling is suddenly switched on, while the latter model produced a much better fit. Saro et al. (2010) compared the galaxy populations in a massive galaxy cluster predicted by Smoothed Particle Hydrodynamics (SPH) simulations with those from the SAM described in De Lucia & Blaizot (2007). Both models included gas cooling and a simple prescription for star formation (as we will do in this study). The resultant object-by-object comparison revealed important differences between the two methods. In particular, the star formation history of brightest cluster galaxies in the SPH simulations is characterized by a more prominent high-redshift peak and lower level of recent star formation with respect to predictions from the SAM. As noticed by Saro

et al. (2010), this is due to the assumption of an isothermal gas density distribution for the hot gas in the SAM, which differs from the actual gas distribution in the simulation.

In De Lucia et al. (2010, hereafter Paper I) we compared the results of stripped-down versions of three independently developed SAMs, 'Durham' (Cole et al. 2000; Benson et al. 2001), 'Munich' (De Lucia & Blaizot 2007; Saro et al. 2010) and 'MORGANA' (Monaco et al. 2007). We ran them on the same set of merger trees taken from the Millennium simulation (Springel et al. 2005). We concentrated on two mass scales, selecting 100 haloes as massive as the Milky Way at $z = 0$ and 100 as abundant as 'SCUBA' galaxies at $z \sim 2$. We found that the resulting cold masses and cooling rates were in good agreement at the Milky Way mass scale, but we noticed that the 'Durham' model predicted systematically lower cooling rates in haloes of the SCUBA set. We showed that this difference is again due to different assumptions on the gas profile, which is isothermal in the 'Munich' model and cored in the 'Durham' model. A comparison of the results of the same models in a configuration where star formation and stellar feedback is active was presented in Fontanot et al. (2013).

More recently, Lu et al. (2011) compared the results of several implementations of semi-analytic cooling models with 1D simulations of both an isolated halo and an accreting halo, and with a simulation of a cosmological volume including inefficient stellar feedback. They reported that different models give predictions of gas accretion rates that can vary by up to a factor of 5. With respect to simulations, SAMs underpredict gas accretion rates in small haloes and overpredict them in massive haloes. They also showed that the predicted cooling-dominated and infall-dominated regimes do not closely correspond to the regimes where cold flows or hot atmospheres were found to dominate in their simulation. A similar setting was used in Hirschmann et al. (2012), with the aim of comparing the predictions of simulations and SAMs when feedback and star formation are used. They used a set of resimulations of DM haloes (see Oser et al. 2010) run with the GADGET code, but they implemented primordial cooling and inefficient stellar feedback, and compared the results with several versions of the model of Somerville et al. (2008), including a stripped-down one with no stellar feedback. This work shows that, while stripped-down SAMs and simulations agree rather well (with some discrepancies consistent with those found in the papers cited above), SAMs predict that the deposition of cold gas takes place mostly in the cooling-dominated regime, while the contribution from simulated cold flows is significant at all redshifts and dominant at $z \gtrsim 1-2$.

To perform an accurate comparison of simulations with SAMs, the SAMs should be run on merger trees extracted from the same simulation, as was done in Yoshida et al. (2002), Helly et al. (2003), Cattaneo et al. (2007), Saro et al. (2010), and Hirschmann et al. (2012) but not in Benson et al. (2001) and Lu et al. (2011). Proper time sampling of merger trees is also relevant: Benson et al. (2012) demonstrated that a SAM gives a convergent description of the formation of galaxies if merger trees are sampled at least ~ 128 times, a factor of 2 higher than, e.g., the time sampling of the Millennium simulation. Finally, the condensation of gas into the central 'galaxy' is governed not only by cooling but also by the time required for progenitor haloes to merge with the central object. As demonstrated in Paper I, different SAMs do not use consistent predictions of galaxy merging times, and this should be properly taken into account.

A further important issue is related to the runaway nature of cooling of self-gravitating gas. When no feedback from stars is present, gas (over-)cools in the small DM haloes that form at high redshift, so cooling takes place in poorly resolved haloes, and in the

infall-dominated regime where the cooling models we are aiming to test do not apply. Increasing the resolution would imply resolving smaller haloes at higher redshift, thus, worsening the overcooling problem. Results showing a reasonable agreement of SAMs and simulations at $z = 0$ may simply reflect the fact that both predict that most baryons have (over-)cooled.

We are convinced that a satisfactory numerical test of semi-analytic cooling models is still missing. An ideal test should have the following characteristics: (i) the study should be limited to cooling in well-resolved haloes; (ii) a clear assessment of the accumulation of cold gas in the cooling-dominated regime, independent of (over-) cooling at high redshift, should be performed; (iii) the analysis should include several SAMs; (iv) SAMs should be run on merger trees extracted from the same simulation; (v) time sampling should be fine enough to guarantee convergence of SAM predictions; (vi) measured cooling rates should be independent of the accuracy with which merger times are predicted by SAMs. In this paper, we present a comparison of SAMs and numerical (SPH) simulations that meets all the above criteria.

Finding a numerical solution of cooling in cosmological haloes is not a straightforward task. To face this problem, we use the `GADGET Tree-PM+SPH` code. The specific implementation used in this paper is shortly described in Section 2.1. While this project was in progress, Kereš et al. (2012) published a comparison of `GADGET` with the new moving-mesh code `AREPO` (Springel 2010), run on the same initial conditions and with a setting similar to that of Lu et al. (2011), though at higher resolution. They reported that galaxy stellar mass functions obtained with the two hydrodynamic schemes tend to be quite similar in the low-mass end but exhibit significant differences in the high-mass end. They traced the origin of this difference to the different efficiencies of dissipative heating from gas accretion on to haloes, which cause cooling to be partially offset in SPH with respect to the moving-mesh hydrodynamic scheme. The issue was readdressed by Nelson et al. (2013), who showed that the structure of cold filaments penetrating into hot haloes is very different for the two hydro solvers. While a detailed comparison of different hydrodynamical schemes is beyond the objectives of this paper, we stress that some of our results might be affected by the use of SPH. We will comment on this below.

As in Paper I, we use here three stripped-down SAMs. The model of De Lucia & Blaizot (2007), denoted the *Munich* model in Paper I, will be called here *delucia*, while *morgana* of Monaco et al. (2007) will retain its acronym. In place of the *Durham* model, we use the highly modular *galacticus* model of Benson (2012) in a configuration that closely resembles that of Bower et al. (2006). We run the three SAMs on merger trees extracted from a collisionless N -body simulation, and compare them with the results of a simulation run on the same initial conditions but including gas hydrodynamics and radiative cooling for a primordial composition. We set merger times to zero in SAMs, and compute cooled masses in simulations by summing over all substructures in simulated friends-of-friends (FoF) haloes. We run the hydrodynamical simulation in two configurations: we allow cooling to be active from the start, or we run the simulation without radiative cooling down to redshift $z \sim 2$, and then switch cooling on. This second configuration allows us to test cooling models exactly in the range of redshift and halo mass where their approximations are expected to be valid: when cooling is switched on, no (over-)cooling has taken place in the infall-dominated regime and all the baryons associated with haloes are in hot atmospheres; the contribution of cooling from small haloes (that are poorly resolved and where gas mostly cools in the infall-dominated regime) is much smaller.

We use a relatively small box ($36 h^{-1}$ Mpc on a side, 50 Mpc for $h = 0.72$) sampled with 512^3 DM particles and an equal number of gas particles, and use a force resolution of 1.5 comoving h^{-1} kpc. With these choices, we prioritize resolution over statistics and test for the first time cooling at a resolution that is sufficient to resolve the morphology of a galaxy when star formation and feedback are properly taken into account (e.g. Scannapieco et al. 2012). The DM halo masses we test range from small galaxies ($3 \times 10^{11} M_{\odot}$) to rich galaxy groups ($5 \times 10^{13} M_{\odot}$). Merger trees have been sampled at 128 time-steps (uniformly in the log of scalefactor; Benson et al. 2012), with the time interval between two snapshots roughly corresponding to half the dynamical time of DM haloes. Moreover, following Viola et al. (2008) and Saro et al. (2010), we use a ‘star formation’ algorithm to remove cooled gas particles from hydrodynamics, with the result of speeding up the simulation and strongly reducing numerical artefacts at the interface of cold and hot phases.

We will be mainly concerned with cooling rates and cooled masses, neglecting for the time being whether gas has cooled on the main substructure of the halo or on a satellite. Other important issues, like the density profiles of gas in cooling haloes, the quantification of cooling on non-central galaxies, or the delineation between cold-flow and hot-flow modes, will be addressed in future works. This paper is organized as follows. Section 2 describes the simulations run for this project and the post-processing analysis used to obtain cooling rates and cooled masses. Section 3 presents the three SAMs used in the paper. Section 4 reports the comparison of results from simulations and SAMs. Finally, Section 5 compares the results with those presented in previous papers, and Section 6 gives a summary of the main results and a discussion.

2 SIMULATIONS

2.1 The code

The simulations in this paper were performed with version 3 of the `GADGET` code (Springel 2005), a massively parallel TreePM+SPH code with fully adaptive time-step integration. To describe the hydrodynamical evolution of gas, the code adopts the SPH formulation of `GADGET2` where entropy is explicitly conserved (Springel 2005). Whenever a Lagrangian SPH code is used, cooling of hot gas at the interface with a cold condensation is affected by a numerical bias, due to the fact that density estimates of hot particles are affected by the many cold SPH neighbours. As shown by Yoshida et al. (2002) and Tornatore et al. (2003), the entropy-conserving formulation of SPH limits the importance of this numerical effect. Another way to limit it is to avoid the formation of cold clumps using a ‘star formation’ recipe to transform cooled particles in collisionless ones (Viola et al. 2008; Saro et al. 2010); this is described below. Radiative cooling of a plasma with primordial composition is computed as in Katz, Weinberg, & Hernquist (1996), assuming an ionizing UV background switched on at $z \sim 6$ and evolving with redshift as suggested in, e.g., Haardt & Madau (1996).

In the absence of heating from star formation, following the evolution of gas subject to runaway cooling is computationally expensive, as most time is devoted to integrating the hydrodynamics of overcooled and unresolved gas condensations with sizes that are much smaller than the gravitational softening. Since cooling is irreversible in these conditions, it is convenient to transform cooled particles into collisionless ‘stars’. We thus used a ‘star formation’ algorithm where every gas particle with overdensity higher than δ_{cold} (with respect to the average gas density) and temperature lower

than $T_{\text{cold}} = 10^5$ K is instantaneously transformed into a collisionless ‘star’ particle. This causes a dramatic speed-up of the code; moreover, the absence of overcooled gas condensations helps in reducing numerical overcooling at the interface of cold and hot gas. This choice has modest drawbacks: the lack of pressure from the cold gas will influence the profile of the hot particles in some way (but in any case the pressure of overcooled gas blobs is unphysical), and collisionless particles will slowly scatter into a ‘diffuse light’ component.

We stress that this ‘star formation’ prescription is not meant to describe the formation of stars but it only represents a convenient numerical shortcut to treat particles that have undergone runaway cooling. In the following, we will refer to ‘star’ particles as cooled particles.

Using a small $9 h^{-1}$ Mpc box with 128^3 DM and 128^3 gas particles and with the same mass and force resolution as the simulations presented here, we performed four test simulations using values of $\delta_{\text{cold}} = 1000, 10000, 30000$, and an effectively infinite value to reproduce to the case of no star formation. Comparing the results of the simulations at $z = 2.9$ (where the simulation without star formation slowed down dramatically), we checked that the density and temperature profiles of hot gas were reasonably stable for changes in δ_{cold} , with 10000 being a good value to preserve hot gas profiles while achieving significant speed-up. Furthermore, we checked that, for the three runs with star formation, the mass of cooled gas changes by no more than ~ 10 – 20 per cent at all redshifts, higher thresholds giving slightly lower values of cooled mass as expected.

2.2 Runs

Initial conditions were created using the public *N-GENIC* code of V. Springel¹, which uses the Zel’dovich approximation to displace particles from a regular cubic grid. We assumed a Λ CDM cosmological model with $\Omega_0 = 0.24$, $\Omega_\Lambda = 0.76$, $\Omega_b = 0.04$, $H_0 = 72 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\sigma_8 = 0.8$, $n_s = 0.96$. We generated initial conditions for a box of $36 h^{-1}$ Mpc (50 Mpc) sampled with 512^3 DM and 512^3 gas particles. The initial redshift was set equal to 199, where the rms particle displacement for the large box is 30 per cent of the grid spacing. Particle masses are $1.93 \times 10^7 h^{-1} M_\odot$ for DM and $3.86 \times 10^6 h^{-1} M_\odot$ for gas. With this choice, a $10^{12} h^{-1} M_\odot$ halo is resolved with $\sim 50\,000$ DM particles. The assumed value for the gravitational softening is $1.5 \text{ comoving } h^{-1} \text{ kpc}$. These values were found in Viola et al. (2008) to be adequate to fully resolve the cooling region of a $10^{12} h^{-1} M_\odot$ halo.

We run this simulation in four configurations. (i) PURED: a pure *N*-body simulation, where all gas particles are treated as collisionless particles. This was used to create merger trees for the SAMs. (ii) NOCOOL: a simulation with gas hydrodynamics but without radiative cooling and UV heating. (iii) COOL: a simulation with radiative cooling, UV heating and ‘star formation’. (iv) COOLZ2: a simulation with the same physics as COOL, using a snapshot of the NOCOOL simulation at $z = 2.089$ as initial conditions. In this paper, we present results based on the PURED, COOL, and COOLZ2 simulations; the NOCOOL simulation will be used in future papers.

To have a proper time sampling of the merger trees (Benson et al. 2012), halo and substructure finding were performed 128 times, uniformly spaced in the logarithm of the scalefactor from $a = 0.05$

to 1. In this way, a halo is sampled roughly two times per dynamical time. This ‘post-processing’ was performed on-the-fly and included a standard FoF halo finder algorithm as well as the substructure finder algorithm *SUBFIND* (Springel, Yoshida, & White 2001). To limit storage requirements, snapshots were saved only 64 times. At $z = 0$, there are 16 haloes more massive than $10^{13} h^{-1} M_\odot$ in the PURED simulation, 159 more massive than $10^{12} h^{-1} M_\odot$ and 302 more massive than $M_{\text{min halo}} = 4.6 \times 10^{11} h^{-1} M_\odot$. This mass roughly corresponds to 20 000 DM particles, and we set it as a lower limit for the mass of the main progenitor. The number of haloes that we find is sufficient to span two orders of magnitude in halo mass and perform averages over many haloes.

Haloes in the COOL and COOLZ2 simulations were matched to those of PURED simulation by checking positions and masses. A few ambiguous cases were found whenever pairs of nearby haloes were classified as merged in one simulation and separated in the other; these haloes were removed from the catalogue.

2.3 Merger trees and halo masses

The construction of merger trees is affected by a number of issues described, for instance, in Fakhouri & Ma (2008) or (Tweed et al. 2009, see also Srisawat et al. 2013). Of the three SAMs used here, the *delucia* and *galacticus* models run by default on merger trees based on DM substructures, while *morgana* runs only on FoF-based trees. In Paper I, all models were run on FoF-based trees, so as to carry out a fair comparison; this choice was tested and discussed there. We adopt the same approach in this paper.

We constructed FoF-based merger trees by directly matching pairs of progenitor/descendant FoF haloes in consecutive outputs. We set a minimum mass of 100 particles for the smallest progenitor considered. Halo pairs were matched when they overlapped by more than 50 per cent of particles, with respect to either the progenitor or the descendant. When substructures are neglected, the main issue is the splitting of haloes, i.e. the case of FoF haloes having two or more descendants. This can be due to either (i) artificial bridging of two separated haloes in one snapshot, which is a byproduct of the FoF algorithm or (ii) highly eccentric orbits of substructures that lead them temporarily out of their main halo. Assuming that a halo A_i at the output i splits into haloes A_{i+1} and B_{i+1} at $i+1$ (A_{i+1} being the largest descendant), we recognize B_{i+1} as an artificially bridged halo (case i) if it verifies these relatively conservative criteria: (1) both A_{i+1} and B_{i+1} overlap by at least 75 per cent (of their particles) with two separate haloes A_{i-1} and B_{i-1} in output $i-1$; (2) the mass of B_{i+1} is at least 1 per cent of that of A_{i+1} ; (3) B_{i+1} contains at least 300 particles; (4) the mass of B_{i+1} is not lower than 90 per cent of that of B_{i-1} . Splitting is then solved by creating a new B_i halo that takes part of the mass of A_i , descends from B_{i-1} , and is progenitor of B_{i+1} . Because bridged haloes are likely to merge in the future, the risk of adopting too conservative criteria is to anticipate the merging time in some cases; we think that this is more acceptable than attempting to separate haloes that have truly merged. When these criteria are not fulfilled, the haloes are considered to belong to case (ii), which is solved by absorbing the split (smaller) halo B_{i+1} back into the main halo A_{i+1} , then iterating the procedure for all descendants of B_{i+1} .

Idealized merger trees obey two further conditions: halo masses should never decrease with time and, after each merger, the mass of the remnant should be larger than that of the merging haloes. Merger trees extracted from simulations do not follow these rules. In SAMs, the amount of baryons associated with a halo is computed as the halo mass multiplied by the universal baryon fraction; so if the former

¹ <http://www.mpa-garching.mpg.de/gadget/right.html>

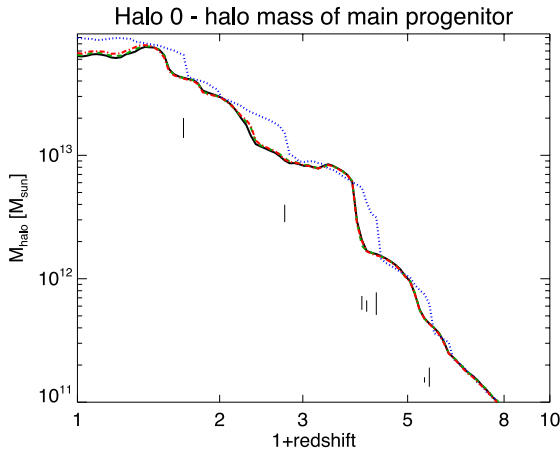


Figure 1. Mass accretion history of the main progenitor of Halo 0, the most massive halo in our box. Black continuous line: M_{200} from the PURED simulation. Blue dotted line: M_{FoF} from the PURED simulation. Green dashed line: M_{200} from the COOL simulation. Red dot-dashed line: M_{200} from the COOLZ2 simulation. Vertical dashes denote major mergers with mass ratios of merging haloes larger than 1:10, the length being proportional to the mass ratio.

decreases, the SAM would lose baryons. In reality, a decrease in halo mass is typically due to a readjustment of the halo profile after a merger, so it is realistic to assume, as we did in [Paper I](#), that whenever halo mass decreases, the baryonic mass available to the halo remains constant. The amount of baryons is thus determined as the cosmic baryon fraction times the largest value of the halo mass reached along the main progenitor branch, including the mass of merging progenitors.

As for the halo mass, we use M_{200} , computed as the mass within a sphere (of radius r_{200}) such that its overdensity is 200 times the critical density of the Universe, $\rho_c = 3H^2/8\pi G$. This quantity is computed by the SUBFIND code, using as centre of the sphere the most bound particle of the main substructure of the FoF halo. Fig. 1 shows the mass accretion history of the main progenitor of the most massive halo in the PURED simulation, hereafter named Halo 0. At $z = 0$, this halo has $M_{200} = 6.49 \times 10^{13} M_\odot$. In this figure, the black continuous and the blue dotted lines correspond to M_{200} and M_{FoF} of the main progenitor in the PURED simulation. The vertical black dashes mark the times corresponding to major mergers, i.e. mergers of haloes with mass ratios larger than 1:10. The length of the dash is proportional to the mass ratio. The two masses grow similarly in time, with the difference that, at major mergers, M_{200} responds more slowly than M_{FoF} to the increase of mass. By visual inspection, we verified that the FoF algorithm merges two haloes when they start to overlap, but the mass of the smaller progenitor significantly overlaps with the r_{200} sphere only towards the end of the relaxation process, which lasts ~ 2 – 3 halo dynamical times. The difference in the two growth curves is a reflection of the fact that there is no obvious definition of ‘mass’ for a halo undergoing a major merger.

Fig. 1 also shows the mass accretion histories (defined via M_{200}) of the same halo in the COOL (green dashed line) and COOLZ2 (red dot-dashed line) simulations. Because cooling concentrates baryonic matter in the halo centre, the M_{200} masses in the cooling runs are slightly larger – $\lesssim 10$ per cent – than in the PURED run. Apart from this, the merger trees are very similar. This degree of similarity is observed in most haloes, with a few exceptions in cases where, at high redshift ($z > 2$), two merging haloes have very

similar masses and the most massive progenitor is not the same in the two simulations. These results confirm that using merger trees from a collisionless simulation induces only small differences in model results.

2.4 Computation of cooled masses

The cooled mass of a halo can be simply computed as the total mass of ‘star’ particles. These particles cool in the deepest regions of potential wells, but then, as noticed in, e.g., Saro et al. (2010), they are subject to tidal forces, if not to two-body heating, and scatter away from the substructure in which they were born to form a ‘diffuse stellar component’. Some of these stars can scatter beyond r_{200} , thus, making the computation of cooled mass more complicated. The cooling rate on a halo can then be computed by taking the difference of cooled masses between two consecutive outputs; in this case, it is necessary to subtract the contribution from gas that cooled in other progenitors and then entered the halo via merging. With post-processing performed on time intervals smaller than the halo dynamical time, we resort to an alternative computation of cooled masses, which avoids the difficulties mentioned above and allows us to easily subtract out the contribution from insufficiently resolved haloes.

GADGET assigns to each cooled particle (‘star’) an ‘age’, equal to the scalefactor at ‘star formation’. Because particles retain their ID number, we identify for each cooled particle the FoF halo it belonged to at the last output where it was still a gas particle. We find that only 0.3 and 0.4 per cent of particles in the COOL and COOLZ2 simulations did not belong to any FoF halo at the last output before cooling. The cooling rate of a halo at a given output is then computed as the mass of all cooled particles, which were gas particles in the previous output and belonged to the halo main progenitor, divided by the time interval between the two outputs. Cooled masses are computed by summing cooling rates over all progenitors along the merger tree.

This approach allows us to subtract out the contribution coming from poorly resolved haloes. In previous work (e.g. Benson et al. 2001; Yoshida et al. 2002; Helly et al. 2003), based on simulations at much poorer resolution, this limit was often estimated by requiring that the halo has at least as many gas particles as the number of neighbours used to compute the density (64 in our case). For our simulation, this would give $\sim 10^9 h^{-1} M_\odot$. This is a very optimistic estimate because cooling takes place in the inner region of the halo, whose size is of order of the scale radius. To quantify the mass scale above which the cooling region is well resolved in our simulations, we use the $9 h^{-1}$ Mpc box already discussed in Section 2.1, run with the same setting as the COOL simulation (identical conclusions are reached using the full COOL and COOLZ2 simulation). We show in the upper panels of Fig. 2, for all progenitors of all haloes more massive than $M_{\text{min halo}}$ and for three redshifts ($z = 0.5, 1$, and 2), the ratio between cooled mass and total gas mass in the halo. At small masses, below roughly $M_{\text{min prog}} = 5 \times 10^{10} M_\odot$, cooled fractions begin to drop to low values. This behaviour is specific to the simulation and is not observed in SAM predictions. This drop is in part an effect of heating due to the UV background. To subtract out this effect, the lower panels report the same quantities for a simulation run *without* a UV background. In this case, cooled masses change significantly below $10^{10} M_\odot$, but the change of trend with halo mass at $M_{\text{min prog}}$ is still present and has to be ascribed to limited numerical resolution.

A halo of mass $M_{\text{min prog}}$ is sampled by $\sim 10\,000$ DM particles, and at $z = 0$, it has a virial radius of ~ 75 kpc; for a concentration of ~ 10 ,

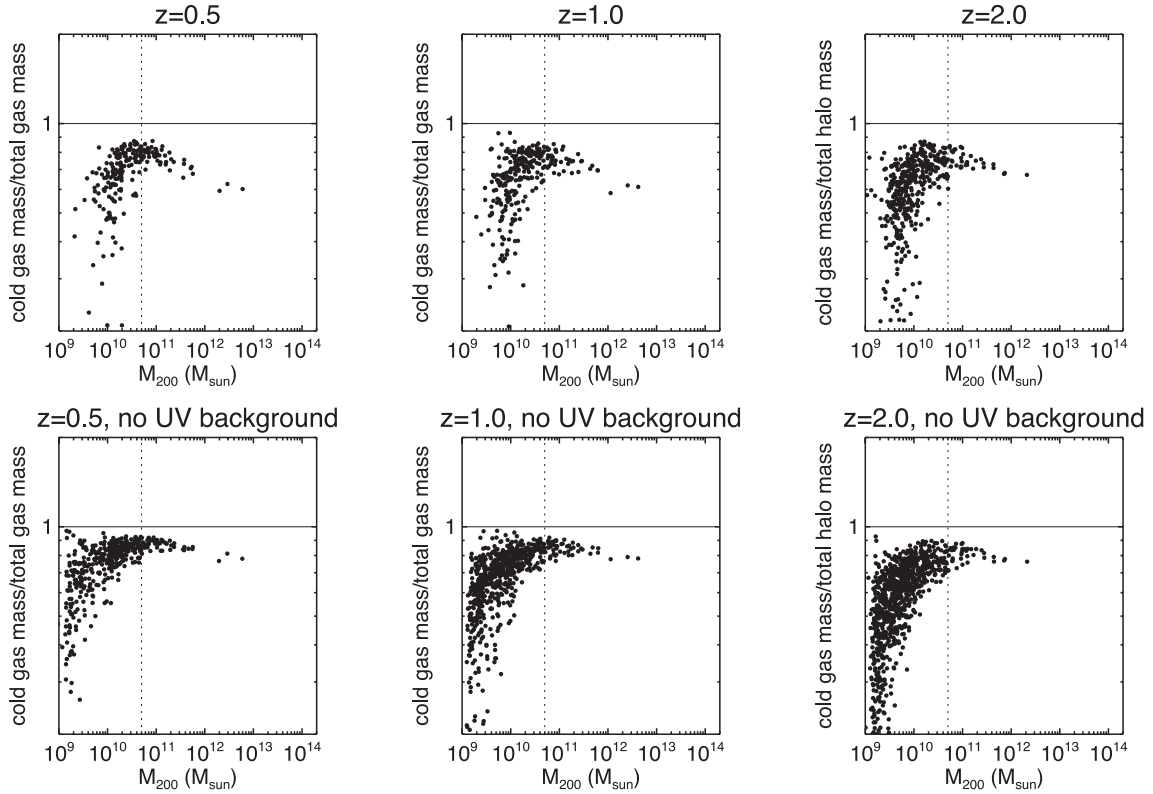


Figure 2. Ratio of cooled and total gas mass versus halo mass for all progenitors of all haloes with $M_{200} > M_{\min \text{ halo}}$, at three redshifts ($z = 0.5, 1$ and 2), for the small $9 h^{-1}$ Mpc box. The vertical dashed line denotes $M_{\min \text{ prog}} = 5 \times 10^{10} M_{\odot}$, the horizontal thin line denotes unity. The lower panels refer to the same simulation run without the UV background.

the scale radius is ~ 7.5 kpc, ~ 5 times the gravitational softening. This is a good, conservative, lower limit for the mass of the smallest halo where cooling is properly resolved in this simulation.

From now on, unless otherwise specified, cooled masses will be computed by summing cooling rates on all progenitors more massive than $M_{\min \text{ prog}}$. In Fig. 3, we show the cooled mass found in the main progenitor of four example haloes that will be used in the following analysis. Two of these are the most massive haloes in the simulation, Halo 0 and Halo 1, with the former suffering a major merger below $z = 1$ and the second having a more quiet merger history at late times. The other two haloes have masses $\sim 10^{12} M_{\odot}$, with Halo 132 suffering a recent major merger and Halo 133 having a quiet mass accretion history. For each halo, we show results from the COOL (dotted blue lines) and COOLZ2 (red dashed lines) simulations. Thick lines give cooled masses computed as described above while thin lines give the total mass of cooled particles in the haloes; accordingly, the differences between the thin and thick lines give the contributions of poorly resolved haloes. This is very significant in all cases for the COOL simulation, amounting to roughly a factor of 2, while it is much more limited for in the COOLZ2 simulation. It is worth reminding that this increase of cooled mass is contributed both by cooling and by mergers.

This figure demonstrates the ability of COOLZ2 to suppress overcooling in poorly resolved progenitors. In the COOL simulation, the fraction of cold mass that has cooled in well-resolved progenitor haloes is roughly constant with redshift (and of order of $\sim 1/2$), because small haloes continually bring cold gas into the main progenitor through mergers. The COOLZ2 simulation produces total cooled masses (the thin lines) that by $z = 0$ have converged to those of the COOL simulation. This is expected, because the fraction

of the residual hot gas will be roughly determined by the density at which cooling times become equal to the Hubble time. But in COOLZ2, most gas cools in haloes with $M_{200} > M_{\min \text{ prog}}$, since high- z overcooling at high redshift is removed and the hot gas density at $z \sim 2$ is higher than in COOL. We thus conclude that in this simulation the numerical description of cooling is less affected by poor resolution than in the COOL run.

The computation of hot masses is affected by an uncertainty related to the definition of halo mass: models use M_{200} to compute the total amount of available baryons, so the output of the post-processing gives the mass of hot gas within the FoF halo. Then, a comparison of models with simulations is affected by the scatter between M_{200} and M_{FoF} . This problem affects the computation of cooled mass described above to a much smaller extent, since particles that suffer runaway cooling are located well within r_{200} . We have verified that the behaviour of hot masses is entirely predictable from that of cooled masses, so for the sake of simplicity we will restrict our analysis to cold masses and cooling rates.

3 SEMI-ANALYTIC MODELS

We compare the results of the simulations described above with predictions from the three stripped-down SAMs *galacticus*, *morgana*, and *delucia* as defined in the Introduction. All models have been adapted to run on the FoF merger trees of the simulation. In all cases, the only baryonic processes implemented are shock heating, cooling, and infall of gas on to the central galaxy. Cooling times are computed using the Sutherland & Dopita (1993) cooling functions assuming zero metallicity. Merger times are set to zero so that galaxy mergers immediately follow the merging of DM haloes.

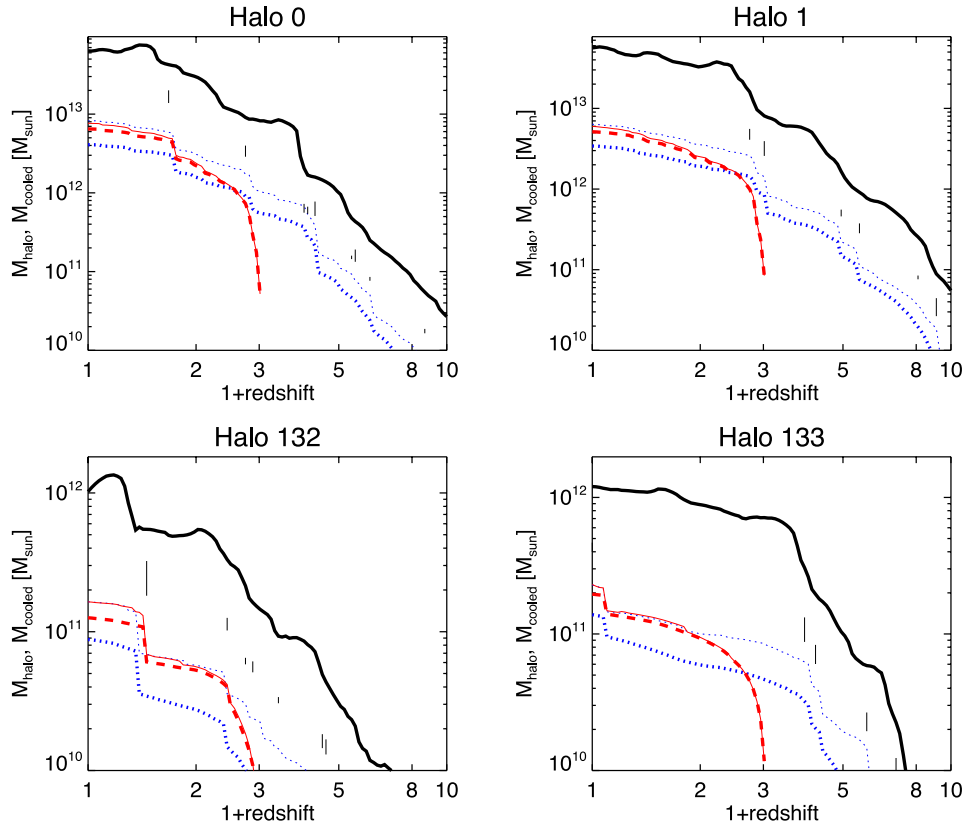


Figure 3. DM and cooled mass accretion histories of the main progenitors of the four example haloes. Black lines give DM masses, with thin vertical dashes marking the position of major mergers. For the coloured lines, thick lines give the mass of cooled gas computed by summing the cooling rates over all progenitors more massive than $M_{\min \text{prog}}$ while thin lines give the total cooled mass. Blue dotted lines refer to the COOL simulation and red dashed lines refer to the COOLZ2 simulation.

Predictions for the COOLZ2 simulation are obtained by switching on gas cooling only after $z = 2.089$. As for the cooling rates, we compute them as the cold mass accumulated on the central galaxy between two snapshots divided by the time interval between them. We verified that these cooling rates are very similar to the instantaneous cooling rates evaluated at the time of the latter snapshot. The cooling and infall models embedded in these SAMs are described at length in the original papers and in [Paper I](#), so we give only a short summary of their main properties here.

3.1 The *delucia* model

The rate of gas cooling is computed following the model originally proposed by White & Frenk (1991), with the specific implementation following that of Springel et al. (2001). The hot gas within DM haloes is assumed to follow an isothermal profile as

$$\rho_g(r) = \frac{M_{\text{hot}}}{4\pi r_{200}^2 r^2}. \quad (1)$$

At each output, the total amount of hot gas available for cooling in each halo is estimated as

$$M_{\text{hot}} = f_b M_{200} - M_{\text{cold}}, \quad (2)$$

where f_b is the universal baryon fraction and M_{cold} is the cold mass associated with the halo. Equation (2) can provide, in a few cases, a negative number (this occurs typically after important halo mergers). In this case, the amount of hot gas is set to zero, and no cooling is allowed in the remnant halo.

The equations for the evolution of gas are solved using 20 time-steps between each pair of simulation snapshots. The cooling time is defined as the ratio of gas specific thermal energy and cooling rate per unit volume:

$$t_{\text{cool}}(r) = \frac{3}{2} \frac{kT\rho_g(r)}{\bar{\mu}m_p n_e^2(r)\Lambda(T, Z)}. \quad (3)$$

In the above equation, $\bar{\mu}m_p$ is the mean particle mass, $n_e(r)$ is the electron density, k is the Boltzmann constant, and $\Lambda(T, Z)$ represents the cooling function for a zero metallicity gas. The virial temperature of the halo is computed as

$$T_{\text{vir}} = \frac{1}{2} \frac{\mu m_H}{k} V_{\text{vir}}^2 \quad \text{or} \quad T_{\text{vir}} = 35.9 (V_{\text{vir}}/\text{km s}^{-1})^2 \text{ K}, \quad (4)$$

where m_H is the mass of the hydrogen atom and μ is the mean molecular mass.

A ‘cooling radius’ is then computed as the radius at which the local cooling time is equal to the halo dynamical time. If the cooling radius is smaller than the virial radius of the halo under consideration, the gas is assumed to cool quasi-statically, and the cooling rate is modelled by a simple inflow equation as

$$\frac{dM_{\text{cool}}}{dt} = 4\pi\rho_g(r_{\text{cool}})r_{\text{cool}}^2 \frac{dr_{\text{cool}}}{dt}, \quad \text{if } r_{\text{cool}} < r_{\text{vir}}. \quad (5)$$

At early times, and for low-mass haloes, the formal cooling radius can be much larger than the virial radius. In this infall-dominated regime, the infalling gas is assumed not to reach hydrostatic

equilibrium and the accretion rate on the galaxy is determined by the halo dynamical time as

$$\dot{M}_{\text{cool}} = \dot{M}_{\text{hot}}/\tau_{\text{dyn}}, \quad \text{if } r_{\text{cool}} \geq r_{\text{vir}}, \quad (6)$$

where $\tau_{\text{dyn}} = R_{\text{vir}}/V_{\text{vir}}$.

3.2 The *galacticus* model

The highly modular *galacticus* code has been configured so as to reproduce the behaviour of the Durham model of Cole et al. (2000), Benson et al. (2003), and Bower et al. (2006).

The hot gaseous component in DM haloes is assumed to have a density profile described by the β -model as

$$\rho_g(r) = \frac{\rho_0}{[1 + (r/r_{\text{core}})^2]^{3\beta/2}}, \quad (7)$$

where ρ_0 is the density at the centre of the halo, r_{core} is the radius of the core, and β is a parameter that sets the slope of the profile at $r > r_{\text{core}}$. The model assumes $\beta = 2/3$; the standard choice of core radius is $r_{\text{core}} = 0.07r_{200}$. The temperature profile of the gas is assumed to be isothermal at the virial temperature (equation 4).

Evolution is integrated between each pair of simulation snapshots using an adaptive time-step integrator set to maintain a tolerance of 1 part in 10^3 in all evolved quantities. Halo masses are linearly interpolated between their values at successive snapshots when needed as input to the integrator.

When the mass of a halo is increasing with time, $\dot{M}_{200} > 0$, the hot halo gains mass at a rate

$$\dot{M}_{\text{hot}} = f_b \dot{M}_{200}. \quad (8)$$

In the cases where the halo is decreasing in mass with time, $\dot{M}_{200} < 0$, this mass-loss is instead accumulated to a quantity, M_{loss} , as

$$\dot{M}_{\text{loss}} = -f_b \dot{M}_{200}, \quad (9)$$

while M_{hot} remains constant (or decreases if cooling is occurring). Once the halo begins to grow once more, the accretion rate on to the halo becomes

$$\dot{M}_{\text{hot}} = \left(f_b + \frac{M_{\text{loss}}}{M_{200}}\right) \dot{M}_{200}, \quad (10)$$

$$\dot{M}_{\text{loss}} = \left(-\frac{M_{\text{loss}}}{M_{200}}\right) \dot{M}_{200}, \quad (11)$$

such that the mass of hot gas does not begin to significantly grow until this halo exceeds the previous maximum value of M_{200} .

The cooling time is defined using equation (3). The cooling radius is computed by equating the radial-dependent cooling time with the dynamical time of the halo. The cooling rate is then given by equation (5) if $r_{\text{cool}} < r_{\text{vir}}$ or by equation (6) if $r_{\text{cool}} \geq r_{\text{vir}}$. The mass of the hot halo is decreased at this rate, and \dot{M}_{cool} is integrated between each pair of simulation snapshots to find the mass of gas that flows on to the central galaxy.

3.3 The *morgana* model

The cooling model implemented in *morgana* is described in Monaco et al. (2007) and Viola et al. (2008). The hot halo phase is assumed to be in hydrostatic equilibrium in an Navarro, Frenk & White (NFW, Navarro, Frenk & White, 1997) halo, to fill the volume between the cooling radius and the virial radius of the halo, and to be subject

to a polytropic equation of state with index $\gamma_p = 1.15$. Under these assumptions, one obtains

$$\rho_g(r) = \rho_{g0} \left(1 - a \left(1 - \frac{\ln(1 + c_{\text{nfw}} x)}{c_{\text{nfw}}}\right)\right)^{1/(\gamma_p - 1)}, \quad (12)$$

$$T_g(r) = T_{g0} \left(1 - a \left(1 - \frac{\ln(1 + c_{\text{nfw}} x)}{c_{\text{nfw}}}\right)\right), \quad (13)$$

where a is a function of halo concentration ($c_{\text{nfw}} = r_{\text{halo}}/r_s$), $x = r/r_s$, and r_s is the halo scale radius. ρ_{g0} and T_{g0} are the values of density and temperature extrapolated to $r = 0$, while $\eta(r) = T_g(r)/T_{\text{vir}}$ and $\eta_0 = \eta(0)$.

The cooling rate is computed by integrating the contribution to cooling from all mass shells as

$$\frac{dM_{\text{cool}}}{dt} = \frac{4\pi r_s^3 \rho_{g0}}{t_{\text{cool},0}} \quad (14)$$

$$\times \int_{r_{\text{cool}}/r_s}^{c_{\text{nfw}}} \left[1 - a \left(1 - \frac{\ln(1+t)}{t}\right)\right]^{2/(\gamma_p - 1)} t^2 dt, \quad (15)$$

where $t_{\text{cool},0}$ is computed using the central density ρ_{g0} and the temperature of the gas at r_{cool} .

At each output, the mass of gas that falls on to the halo is computed as

$$M_{\text{infall}} = f_b \cdot \max \left[M_{200} - \left(\sum_i M_{200}^i + M_{200}^{\text{max}} \right), 0.0 \right], \quad (16)$$

where the sum extends over all progenitors of the halo other than the main one and M_{200}^{max} is the maximum value of the virial mass along the main progenitor branch. The evolution of the system is followed with a Runge–Kutta integrator with adaptive time-steps, and the equilibrium configuration of the hot halo gas is recomputed at each time-step.

By equating the mass cooled in a time interval dt with the mass contained in a shell dr , one obtains the evolution of the cooling radius, which is treated as a dynamical variable in this model, as

$$\frac{dr_{\text{cool}}}{dt} = \frac{dM_{\text{cool}}/dt}{4\pi\rho_g(r_{\text{cool}})r_{\text{cool}}^2} - c_s \quad (17)$$

where c_s is the sound speed computed at r_{cool} , and is added to the right-hand side of the above equation to allow the ‘cooling hole’ to close at the sound speed.

The calculation of cooling rate is started when a halo appears for the first time, with $r_{\text{cool}} = 0.001r_s$, and is reset after each halo major merger. Cooling is then assumed to be quenched for n_{quench} dynamical times after each merger, plus one cooling time, to mimic the time needed by the merging system to relax into the new configuration and start cooling again. We use the standard value of $n_{\text{quench}} = 0.3$.

4 RESULTS

As in Paper I, we illustrate the behaviour of models both by considering the predicted cooled masses and cooling rates for all haloes more massive than $M_{\text{min halo}}$ at some relevant redshifts and by showing detailed predictions for the four example haloes selected in Section 2.4. We recall that Haloes 0 and 132 experience major mergers below $z = 1$, while Haloes 1 and 133 have more quiet merging histories. To present a more organic discussion, we show all the main results before discussing them.

Figs 4 (COOL simulation) and 5 (COOLZ2 simulation) show, for each halo or its main progenitor, a comparison between model predictions and simulation results. Starting from the upper panels, we give results at $z = 0$ for the three models separately; *galacticus* is shown in red, *morgana* in blue, *delucia* in green. For each set of 3×2 panels, the upper panels report the cooled masses, the lower panels the cooling rates. Pairs of points corresponding to the same halo are connected by a thin black line; this allows us to highlight the cases in which SAM predictions lie beyond the limits of the plot. In all panels, the lower part reports the ratio between SAM and simulation results, thick lines show median values of the points in mass bins of 0.5 dex, while thin lines correspond to the 16 and 84th percentiles computed as follows: calling R_i the ratio of model and simulation predictions for halo i , we compute for each mass bin the median value $R_{\text{med}}(M)$ and the quantity $x_i = R_i/R_{\text{med}}(M)$, then compute the 16 and 84th percentiles of the set of $\{x_i\}$ for all haloes and show these percentiles scaled with the median value. Results are given at $z = 0, 1, 2$, and 3 for COOL (Fig. 4), and $z = 0$ and 1.5 for COOLZ2 (Fig. 5).

The detailed evolution of the four example haloes is shown in Figs 6 (COOL simulation) and 7 (COOLZ2 simulation). For each halo, we show its mass accretion history for DM and cold gas, cold gas fraction, and cooling rate of the main progenitor. Simulation results are shown in black; colour coding for the models is the same as in the previous figures.

We confirm that cooling models embedded in SAMs are able to follow relatively well the accumulation of cooled mass obtained in the simulation. As far as cooled masses are concerned, average differences are smaller than ~ 30 per cent, with some exceptions that will be commented on later (Figs 4 and 5), and only a few haloes show values that are discrepant by more than a factor of 2. Cooled masses are also recovered with a relatively small scatter, typically less than 20 per cent. Cooling rates are recovered in a much noisier way, with average discrepancies and scatter being typically within a factor of 2 for *galacticus* and *delucia*, within a factor of 3 for *morgana*. The same level of agreement is achieved in the detailed history of example haloes (Figs 6 and 7). Beyond this overall agreement, a number of interesting discrepancies are evident.

(i) Cooled mass and cooling rates predicted by SAMs are typically below the corresponding results from the COOL simulation at $z \gtrsim 2$. This is apparent both in the histories of the four example haloes (Fig. 6) and in the $z = 2$ and 3 panels of Fig. 4, where cooled masses are below the simulated ones by ~ 20 –40 per cent, while cooling rates are in much better agreement. This implies that models are slow in accumulating cooled mass in the infall-dominated regime. This result is evident only when the mass cooled on unresolved haloes is subtracted out. Fig. 8 shows how the cooled mass is recovered by one of the models (*delucia*) when the contribution from poorly resolved haloes is (left) or is not (right) subtracted out both in the model and in the simulation. Model results are in better agreement with simulations in the second case, but this agreement is due to the fact that the numerical underestimate of cooled gas happens to roughly compensate the model underestimate. Similar results are obtained for the other two models.

(ii) For massive haloes, the *galacticus* model underestimates cooled masses and cooling rate relative to simulations. This is clearly visible for Haloes 0 and 1, for which the *galacticus* prediction of cooling rates drops to zero at $z < 1$. Figs 4 and 5 shows that at $z = 0$, cooling rates for the most massive haloes ($M > 2 \times 10^{14} M_\odot$)

are systematically below simulation values and approach 0 in several cases.

(iii) *morgana* reproduces the simulated cooling rates with much more scatter than the other two models: the fraction of haloes discrepant by more than a factor of 10 ranges from 13 to 20 per cent (10 to 15 per cent for the other models). As apparent in Figs 6 and 7, this is connected to the behaviour of the model at major mergers: cooling is quenched for a period that corresponds to n_{quench} dynamical times plus one cooling time, while no such behaviour is seen in simulations. Cooling rates then peak to high values and compensate the underestimate, thus providing on average correct values for the cooled mass, though the scatter with respect to simulated values is a higher than the other models.

(iv) At low redshift, both *delucia* and *morgana* tend to overestimate cooling rates in massive haloes. This is consistent with the finding of Saro et al. (2010) of higher cooling rates at low redshift in a galaxy cluster. In that paper, this discrepancy was shown to be due to the isothermal profile assumed for the gas in the *delucia* model. The same trend is, however, present at very similar level in *morgana*, which assumes a hydrostatic density profile with a shallow inner slope closer to the actual gas profile found in the simulations. We will come back to this issue later.

(v) Switching on cooling at $z = 2$ has the effect of triggering quick deposition of cooled mass and high cooling rates. From Fig. 7, we see that, while *delucia* and especially *morgana* follow this transient well and quickly converge to the cooled mass of the simulation, *galacticus* predicts lower cooling rates and is slower in converging to the correct gas fraction; in massive haloes, it always remains below the simulation value. Looking at Fig. 5, we see that *morgana* gives a good fit of simulated cooled fractions already at $z = 1.5$, though with substantial scatter, where *delucia* is biased low by ~ 30 per cent and *galacticus* by a factor of 2.

4.1 The role of gas profile

Point (ii) is in line with the finding in Paper I that the cored gas profile assumed in the Durham model is mainly responsible for the underprediction of cooling rate with respect to the *delucia* model, which assumes an isothermal gas density profile. In Paper I, we verified that, when a similar assumption is made for the gas profile in the Durham model, results from the two SAMs are much closer. Point (v) is in line with the results of Viola et al. (2008), who found that their implementation of ‘classical cooling’, very similar to the cooling model of *galacticus* for the case of a static halo, underestimates cooling rates when cooling is suddenly switched on. In that case, the gas profile was computed assuming a hot atmosphere in hydrostatic equilibrium (as in *morgana*, equations 12 and 13), and the resulting profile gave a good fit to the simulated one before cooling was switched on.

To understand the influence of the assumption of gas profiles in the cooling model, we run the *galacticus* model with a singular isothermal profile or a cored profile with a much larger core radius, of 0.42 times the virial radius. Figs 9 and 10 show some relevant results. As an example for this case, we chose Halo 1, which has a more quiet merging history. Fig. 9 shows for the COOL and COOLZ2 simulations the evolution of Halo 1 for the three *galacticus* models, Fig. 10 shows cooled masses and cooling rates for all haloes; for the COOL simulation, we show results at $z = 0$, while for COOLZ2, we show results at $z = 1.5$ (recall that cooling was switched on at $z = 2$ in this run). In agreement with results discussed in Paper I, we find that using an isothermal gas density profile brings SAM predictions in much better agreement with results from the COOL

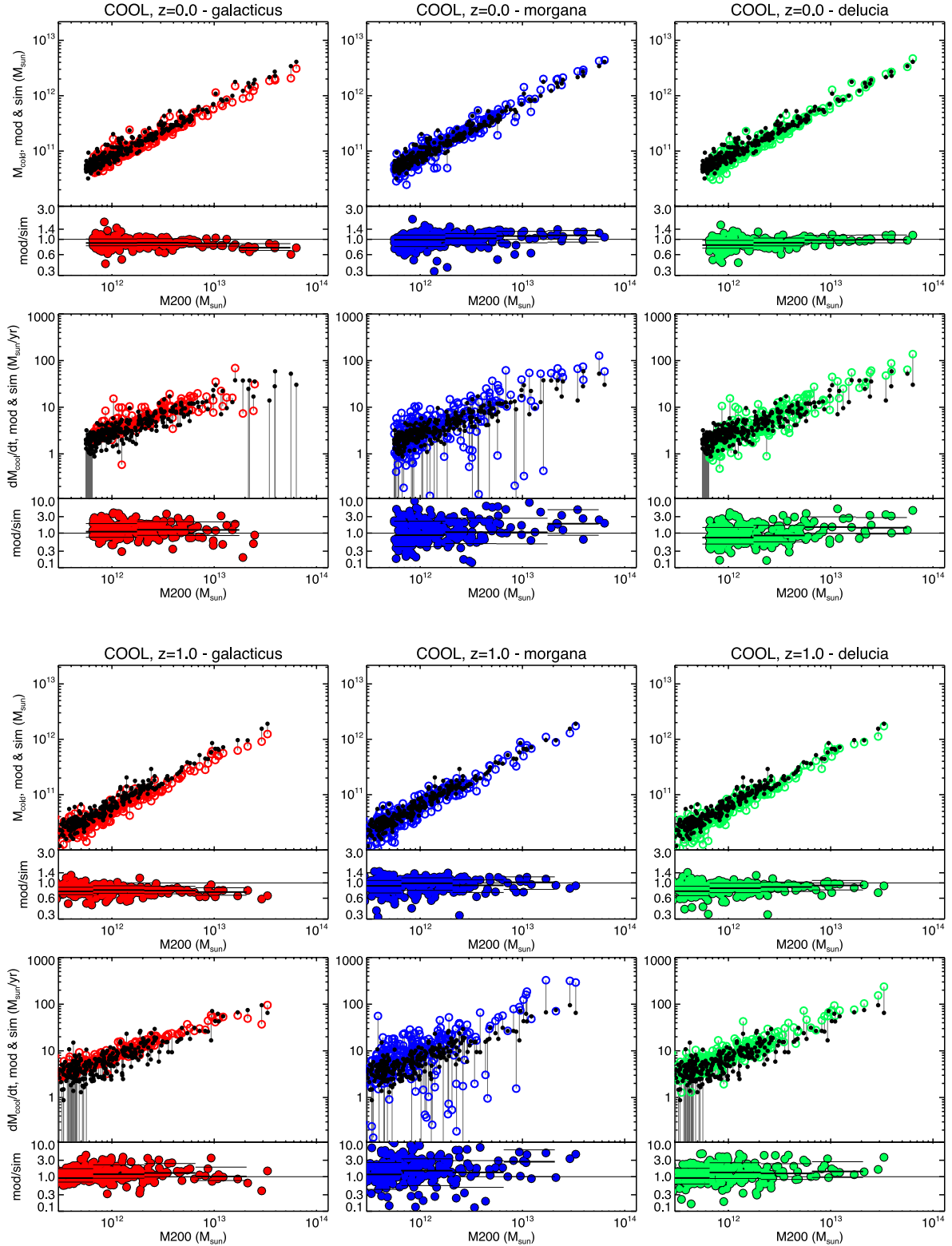
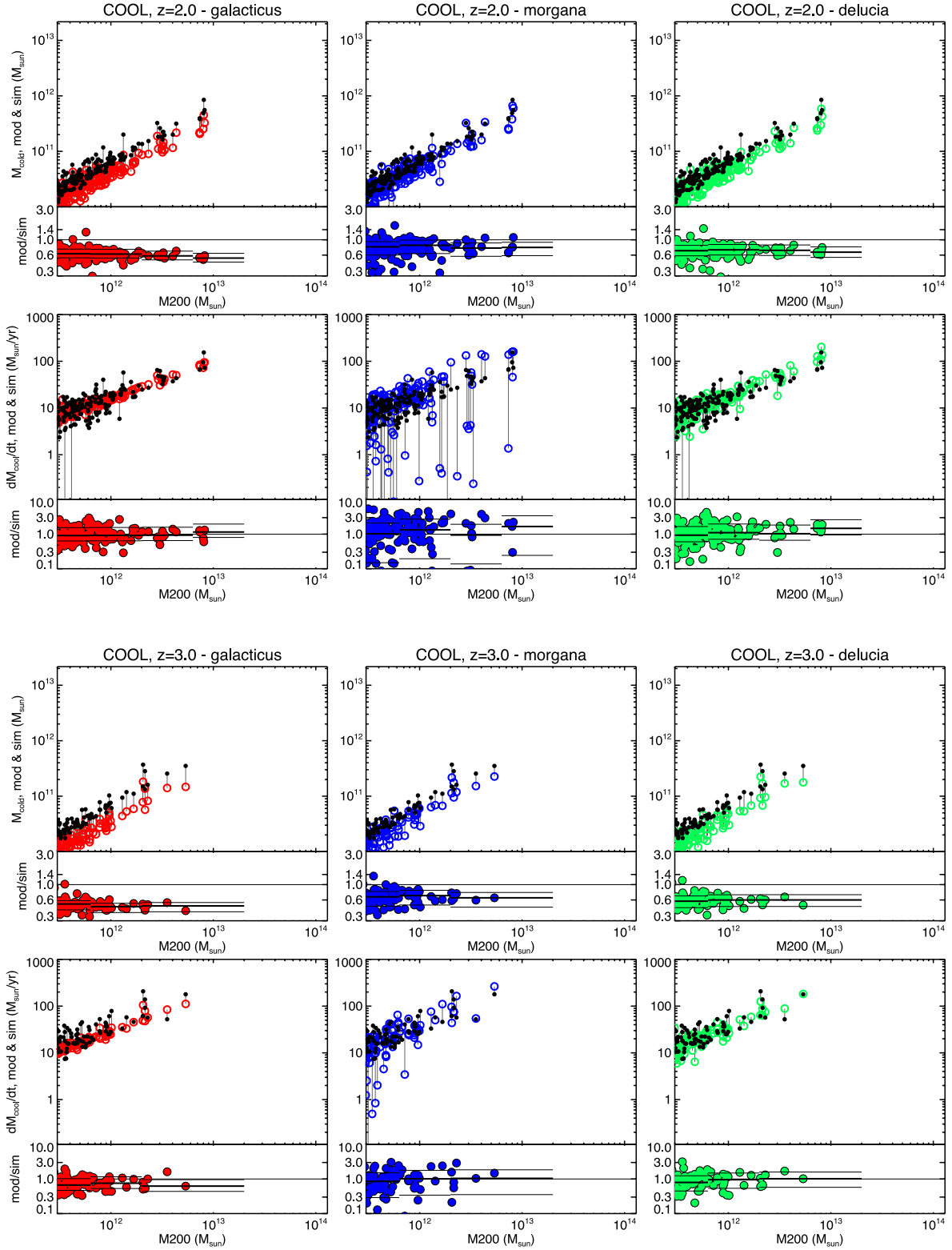


Figure 4. Cooled masses and cooling rates as predicted by models and simulations, for all the haloes with mass higher than $M_{\text{min halo}}$ and their main progenitors. For each panel, the upper part gives, as a function of M_{200} , the cooled mass or the cooling rate as found in the simulations (small black dots) and as predicted by the SAM (open coloured circles). Pairs of points corresponding to the same halo are connected by a thin black line; this allows us to highlight the cases in which SAM predictions lie beyond the limits of the plot. The lower part of each panel gives the ratio of the two quantities, again as a function of M_{200} , while the horizontal lines give medians and 16th and 84th percentiles in halo mass bins. Each group of 3×2 panels reports cooled masses (upper three panels) and cooling rates (lower three panels) for the *galacticus* (right-hand panels, red points), *morgana* (middle panels, blue points), and *delucia* (left-hand panels, green points). From the top, results are given at $z = 0$ and 1, while the second part of the figure gives the same quantities for $z = 2$ and 3.

Figure 4 – *continued*

simulation at low redshift. In this case, significant cooling rates are obtained in large haloes at $z = 0$, while results are very similar for smaller haloes. In fact, this model behaves very similarly to *delucia*, which assumes the same gas density profile; this similarity includes

the tendency of cooling rates higher than simulated ones in massive haloes.

However, when the COOLZ2 simulation is considered, the assumption of an isothermal gas density does not increase significantly

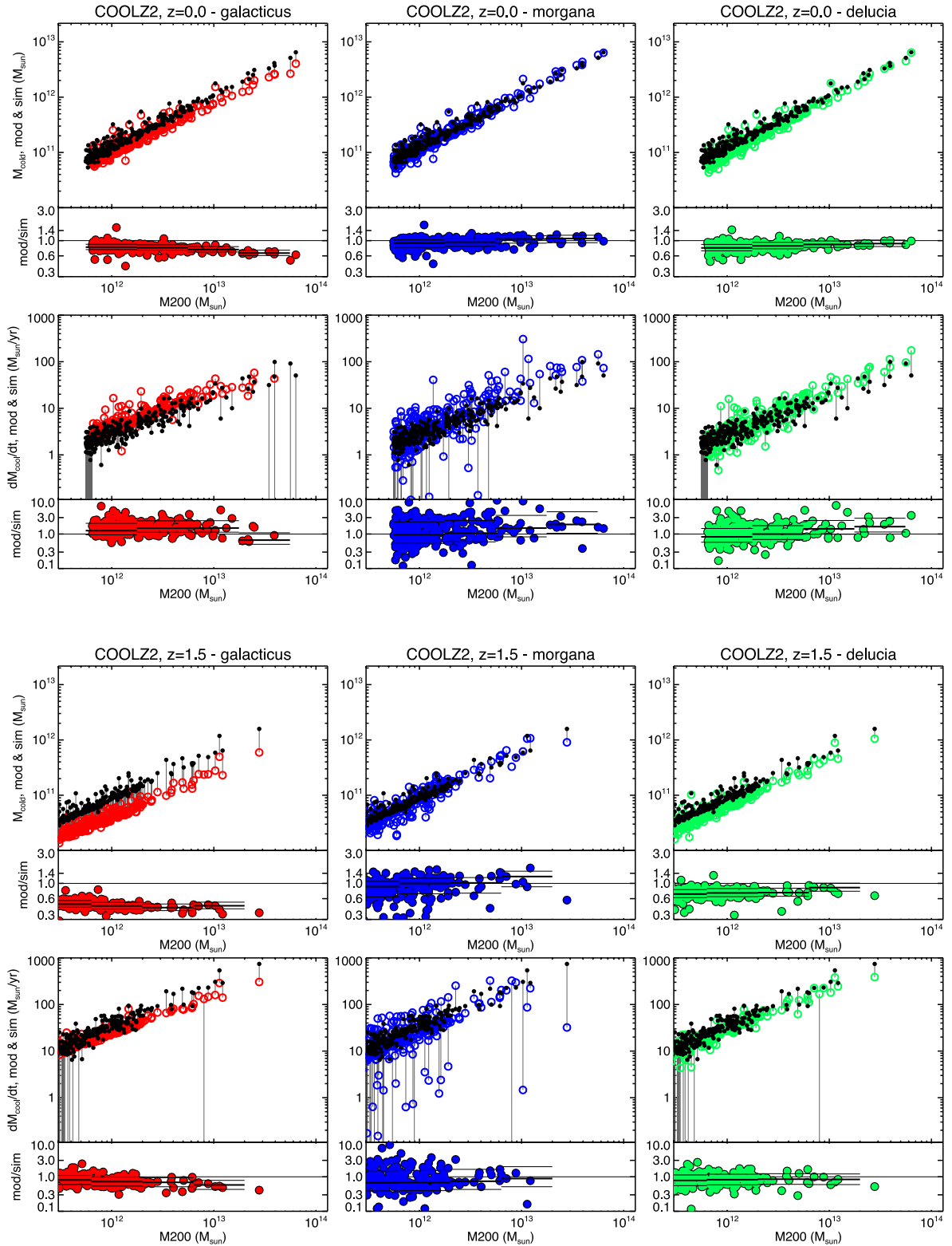


Figure 5. The same as Fig. 4, for the COOLZ2 simulation. From the top, results are given at $z = 0$ and 1.5.

the fraction of cold gas at late times (Fig. 9). As visible in the lower panels of Fig. 10, cooled masses at $z = 1.5$ are as low (50 per cent lower than the simulation) as in the standard core model, while the *delucia* model is lower by 30 per cent only (Fig. 5). This is

in line with Viola et al. (2008), where it was shown that the slow accumulation of cooled mass in the ‘classical model’ is due to the assumption, implicit in equation (5), that each gas shell cools on a cooling time computed using the initial density and temperature,

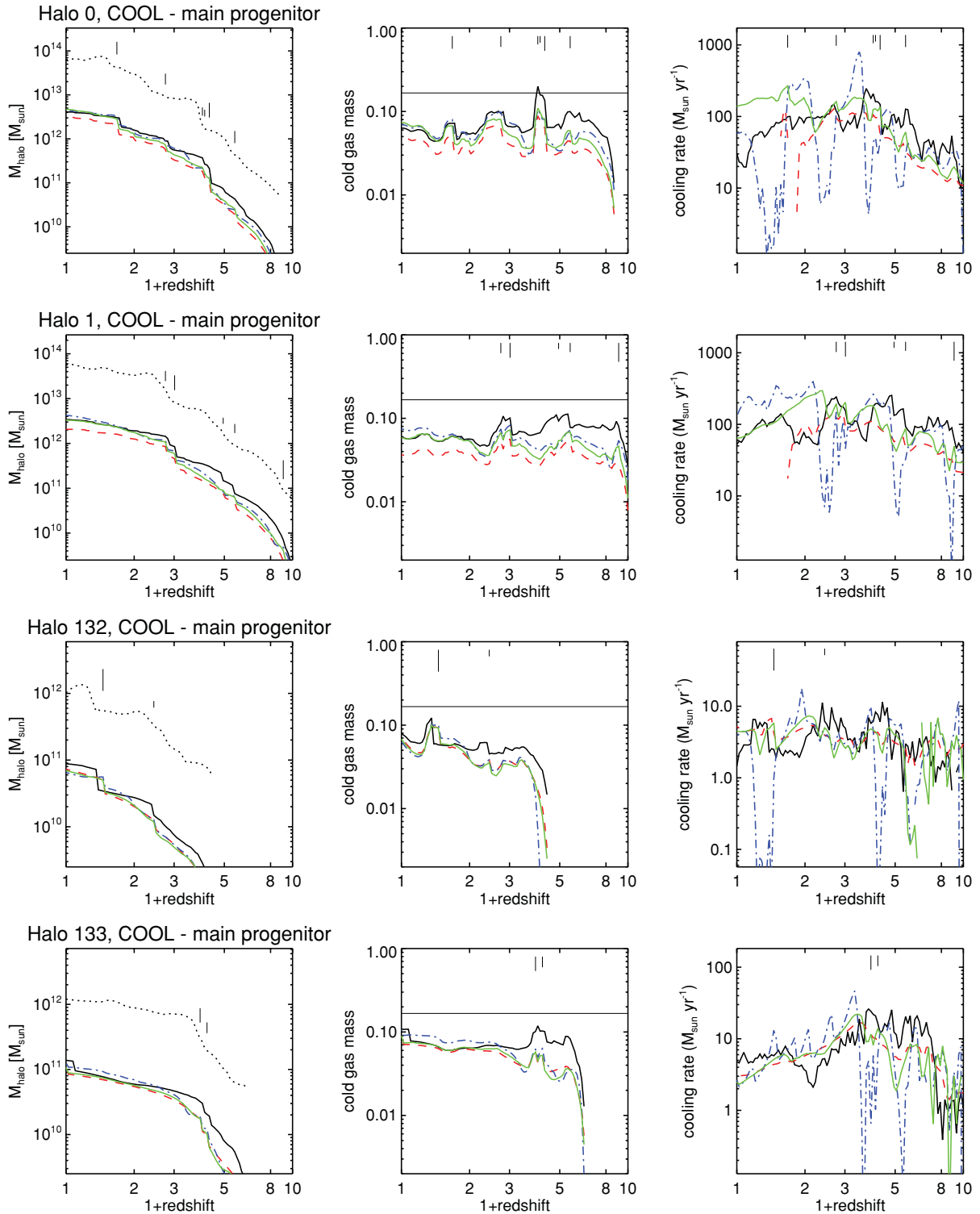


Figure 6. Evolution of cold gas mass, cold mass fraction, and cooling rate of the main progenitor for the four example trees in the COOL simulation. Black continuous lines report the quantities for the simulation, red dashed, blue dot–dashed, and green triple dot–dashed lines give results for the *galacticus*, *morgana*, and *delucia* models. In the left-hand panel, the dotted lines give the value of M_{200} . In the middle panels, the horizontal line denotes the universal baryon fraction. Thin vertical dashes mark major mergers.

while the evolution of gas elements is at increasing density and roughly constant temperature, and this leads to catastrophic cooling on a shorter time-scale. The different results of *delucia* and *galacticus* show that the result is sensitive to the precise integration scheme of equation (5), however.

The results discussed above show that simulations disfavour the use of a very large core for the gas profile in the *galacticus* model, at least in this setting where gas density profile is not affected by any source of feedback. Clearly, feedback is going to influence gas profile and may increase the core size but, whenever the halo is well

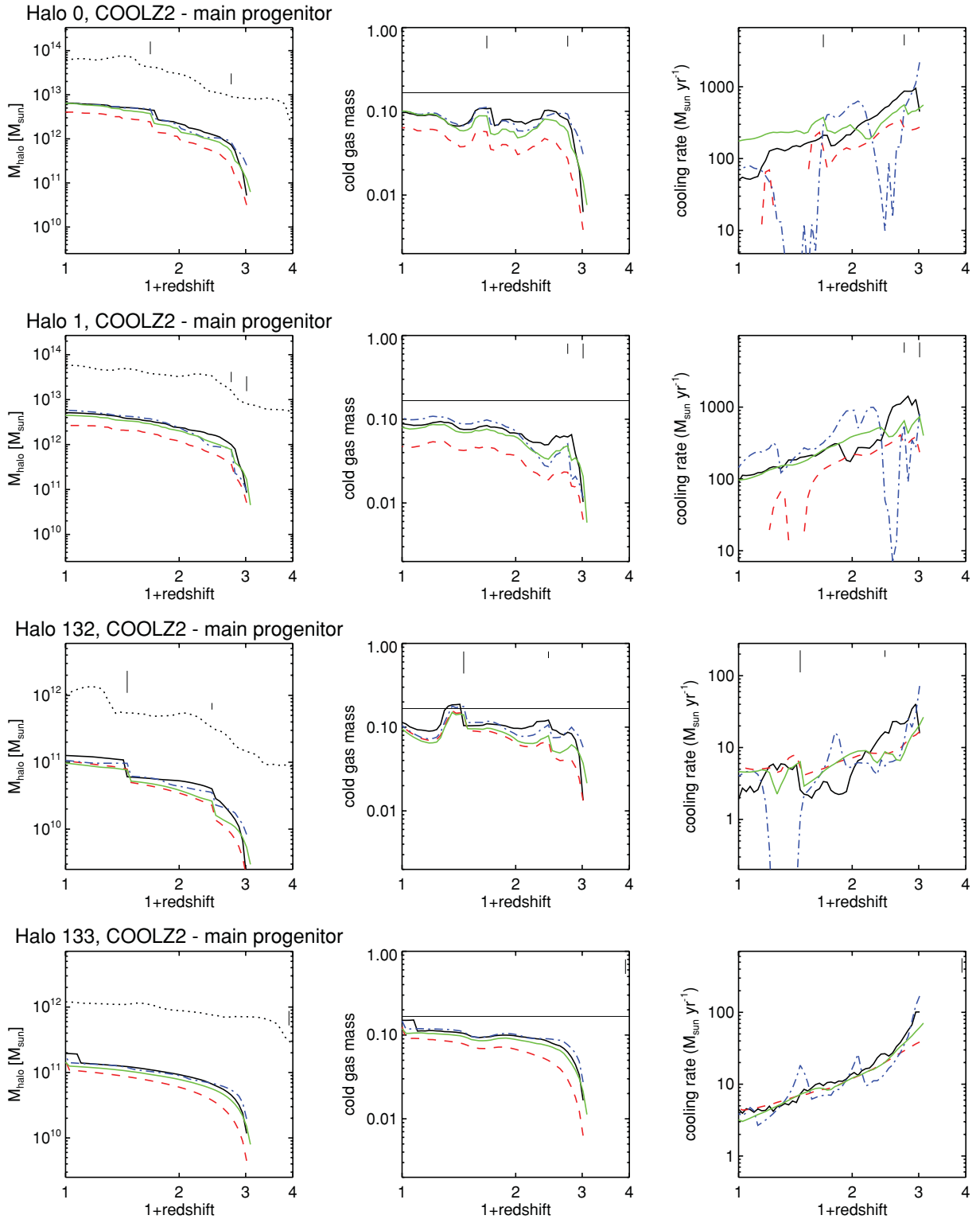


Figure 7. The same as Fig. 6 for the COOLZ2 simulation.

within the cooling-dominated regime, feedback from star formation will be triggered by cooling itself, so it would be difficult to sustain a very large core radius if this has such a large negative impact on the cooling rate.

Another parameter related to the halo density profile is the polytropic index γ_p of the *morgana* cooling model (equations 12 and

13). Because this quantity is constrained both by simulations and, for galaxy clusters, by observations, there is no much freedom to vary it. We ran the model assuming $\gamma_p = 1.05$ and 1.25 , and found that the two new realizations were giving slightly larger discrepancies with respect to the simulations, consistently with 1.15 being the optimal value.

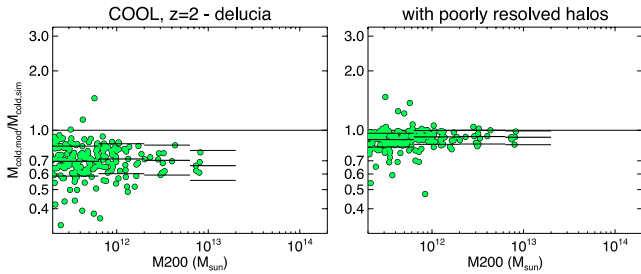


Figure 8. Ratio of cooled mass from the *delucia* model and from the COOL simulation at $z = 0$, as a function of halo mass. Left-hand panel: cooling on poorly resolved progenitors (less massive than $M_{\text{min prog}}$) is subtracted out. Right-hand panel: all progenitors are considered.

4.2 Behavior during major mergers

One reason why *morgana* produces a large scatter in the predicted cooling rates is that it assumes that cooling is switched off in major mergers. The rationale behind this idea is the deep reshuffling of phase space that takes place at these events. Indeed, looking at Figs 6 and 7, no obvious decrease in cooling rates is visible at or after major mergers in simulations. We looked for correlations between the ratio of progenitor masses in two consecutive outputs and the ratio of cooling rates before the merger and after n outputs, with n varying from 1 to a few. The first quantity is large when a major merger takes place; so if major mergers quench cooling, then we would expect the second quantity to be negatively correlated with the first one. We did not find any such correlation.

The lack of any quenching of cooling at major mergers is plausibly connected to the persistence of strong condensations at the

centre of DM haloes. In the case of runaway cooling studied here, the resulting condensations deepen the halo potential well and thus cause an enhancement of hot gas density and then of radiative cooling around them. In other words, they act as seeds for cooling of hot gas. These condensations are simulated as clumps of collisionless stars, and are so compact that large-scale tides are ineffective at disrupting them. The effect of merging-induced shock waves on the hot gas that collect around such condensations will be to some extent affected by the hydro solver, but this should be a further-order effect. It would be very interesting to test this idea with a Eulerian simulation that uses the same setting.

Conversely, when the whole physics of galaxy formation is taken into account, ‘condensations’ will not represent unphysical condensations but true galaxies. In this case, further heating will be provided by energetic feedback from massive stars and AGN. This heating source will be able to limit the accumulation of mass in galaxies and to flatten the density profile of hot gas around them, keeping it to a higher adiabat and making it easier to heat. In these conditions, major mergers may still cause a quenching of cooling. Of course AGN feedback would likely be more effective in this regard.

It is interesting to see what happens to *morgana* when quenching of cooling at major mergers is not applied. Fig. 11 shows, for the COOL simulation, the cooling rates of the two example haloes that suffer major mergers at late times (0 and 132). The continuous line shows the standard cooling model, while the dashed line shows a model with no quenching. Here, the cooling radius is still reset to zero at major mergers; we have verified that dropping this condition does not lead to significant changes. The model without quenching removes some troughs in the cooling rates. In Fig. 12, we show the cooled gas and the cooling rates predicted with the two models

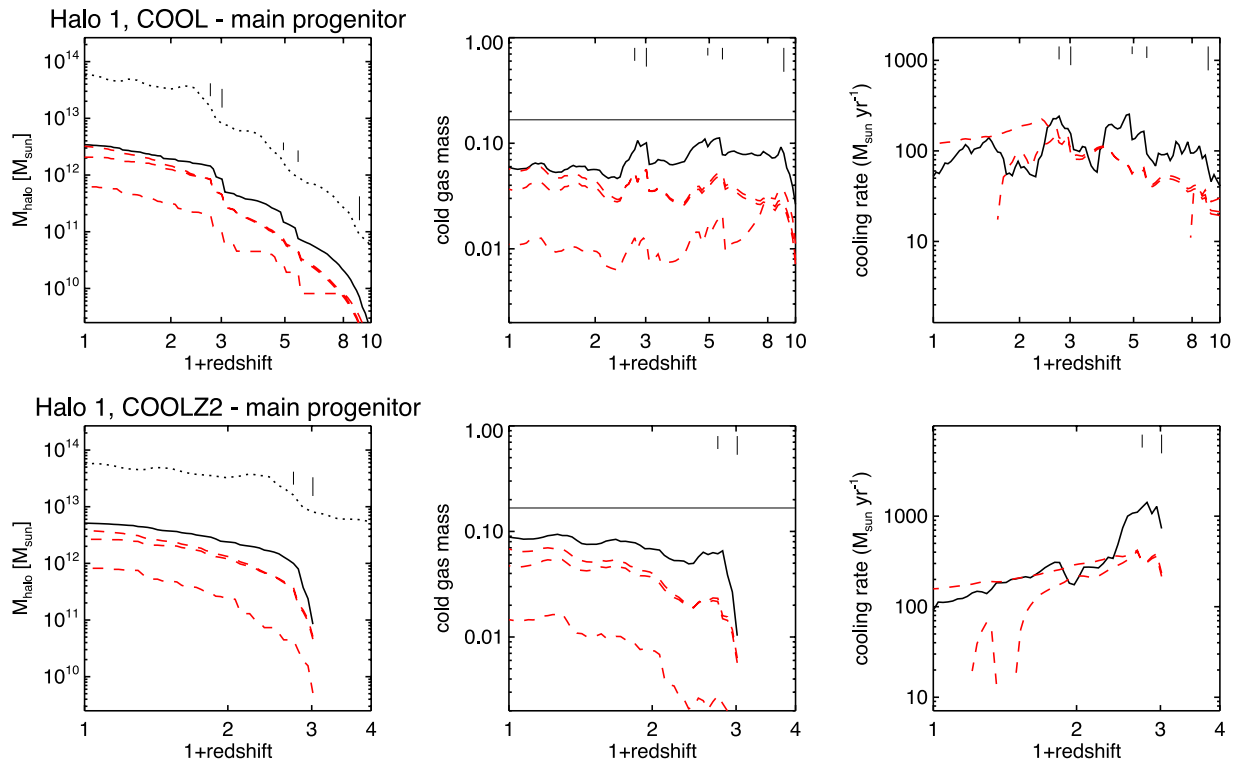


Figure 9. Results of the *galacticus* model with different gas density profiles. The upper and lower rows give (as in Figs 6 and 7), for the COOL and COOLZ2 simulation, the evolution of Halo 1 for standard core (continuous line), isothermal (dotted line), and large core (dashed line) gas profiles.

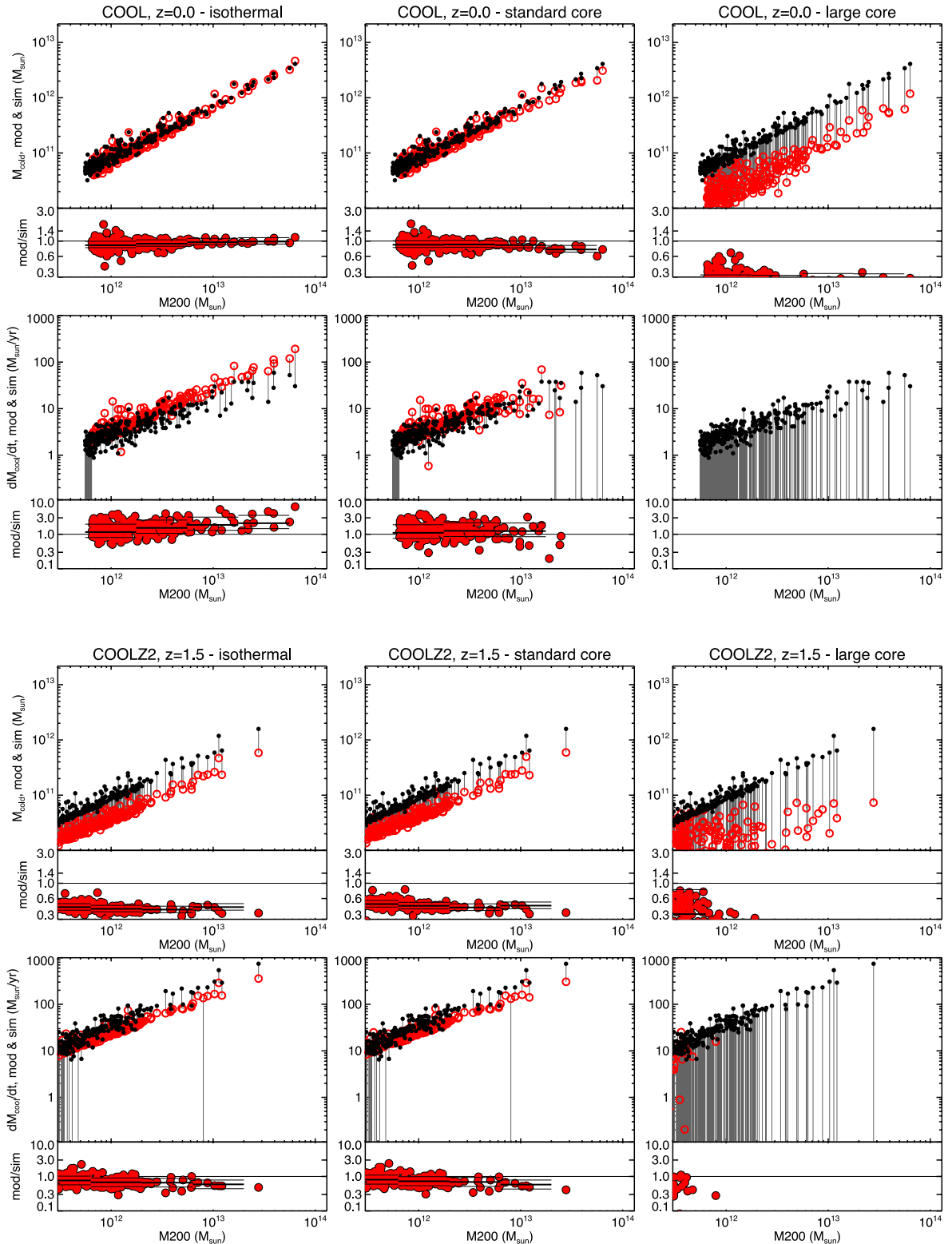


Figure 10. Results of the *galacticus* model with different gas density profiles. The upper and lower panels give (as in Fig. 4) cooled masses and cooling rates for all haloes. Upper panels give the results of the COOL simulation at $z = 0$, lower panels those of the COOLZ2 simulation at $z = 1.5$. Left, middle, and right panels refer to the *galacticus* model with isothermal profile, standard core radius, and large core radius.

for all haloes. Left- and right-hand panels give the results of the standard and no quenching model, respectively. As expected, the number of haloes where cooling rates differ from the simulation values by more than a factor of 10 is less in the no quenching

model, but several cases are still present and the scatter does not decrease significantly. Because of the systematically higher cooling rates, the cooled masses are now biased high by ~ 10 per cent at high masses.

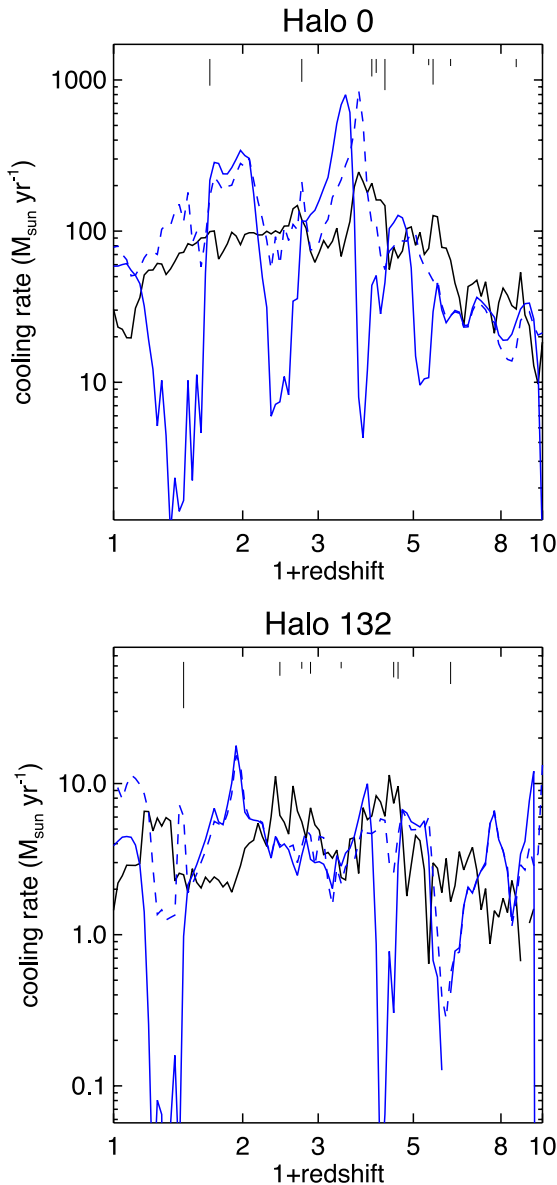


Figure 11. Evolution of cooling rates in the COOL simulation for the two example haloes with late mergers (0 and 132). Results of the *morgana* model are shown for two different assumptions on the behaviour at major mergers. Continuous line: standard cooling model; dashed lines: cooling is not quenched at major mergers.

One more possibility to investigate is to remove the term in equation (17) that decreases the cooling radius at the sound speed. This term is peculiar to the *morgana* cooling model, where the cooling radius is treated as a dynamical variable. This term helps the gas to drift towards the halo centre, so it causes higher cooling rates at late times after a reset of the cooling radius. We have verified that if we drop this term in the model with no quenching we obtain slightly lower cooling rates, which limit the overestimate of cooled mass with respect to the simulation, but the resulting cooled masses are biased low by ~ 30 per cent at low masses.

5 COMPARISON WITH PREVIOUS RESULTS

With respect to previous papers that performed similar tests based on SPH (Benson et al. 2001; Yoshida et al. 2002; Helly et al. 2003;

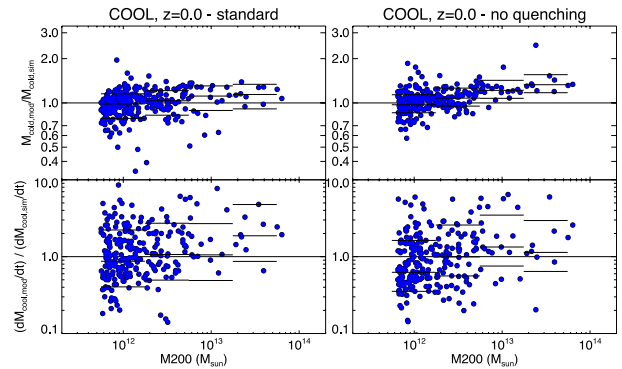


Figure 12. Cooled masses and cooling rates for all haloes in the COOL simulation at $z = 0$. Results of the *morgana* model are shown for two different assumptions on the behaviour at major mergers. Left-hand panels: standard cooling model; right-hand panels: cooling is not quenched at major mergers.

Cattaneo et al. 2007; Lu et al. 2011), our simulations improve significantly in terms of mass and force resolution. Hirschmann et al. (2012) worked at a resolution and in a mass range analogous to the one used in this paper, while Saro et al. (2010) presented results for a galaxy cluster at a resolution that is analogous to ours for that mass scale. Together with our good time sampling, good mass resolution has allowed us to limit the computation of cooled masses to those haloes that are resolved with more than $\sim 10\,000$ particles. Our conservative limit is much higher than what has commonly been used, and guarantees a proper sampling of the central region where cooling takes place. This high value of $M_{\min \text{ prog}}$ also removes the contribution to cooling from those haloes that are affected by the UV background. As an example of the importance of this effect, we showed (Fig. 8) that the underprediction of the amount of cold baryons in the infall-dominated regime (also reported in other papers, see below) can be compensated for by the effect of poor resolution when much cooling takes place at high redshift.

A detailed comparison of our results to the most recent papers (Saro et al. 2010; Lu et al. 2011; Hirschmann et al. 2012) is complicated by the different treatment of merging times and resolution effects. All papers agree in identifying, beyond a general broad agreement, a trend of underestimation of cooled masses at high redshift or small halo mass and overestimation at low redshift or high halo mass, trends that we find also in our analysis. The high-redshift/low-mass trend was explained by Lu et al. (2011) and Hirschmann et al. (2012) as an effect of the more extended dominance of cooling-dominated regime with respect to SPH simulations, where cold flows are found to easily penetrate the hot haloes and quickly deposit their mass into the central galaxy. The low-redshift overestimation was interpreted by Saro et al. (2010) as an effect of their assumption of a singular isothermal profile for the hot gas, which is at variance with the flatter profile found in simulations. To this last conclusion, we can add that the assumption of a cored profile in *galacticus* leads to an underestimation of cooled masses, while the assumption of hydrostatic equilibrium profile in *morgana* leads to a similar overestimation. This shows how the details of the implementation are of great importance.

The only analogue of our COOLZ2 simulation is the simulation of static haloes performed by Viola et al. (2008). Our results are fully consistent with those presented in that paper and they extend the same conclusions in the much more realistic environment of full DM merger trees: following the quick onset of cooling is a challenge for cooling models. The poor performance of a model,

like *galacticus*, based on Cole et al. (2000) was explained in that paper by showing that equation (5) is strictly valid as long as a mass element cools to low temperatures in a time equal to the cooling time computed using its initial density and temperature, but gas elements evolve at roughly fixed temperature and increasing density, and thus cool more quickly. While all models use some version of equation (5) for computing the cooling rate, the specific implementations makes the results very different for this test.

The difference between *GADGET* and the moving-mesh code *AREPO*, reported by Kereš et al. (2012) and Nelson et al. (2013), raises the worry that our results may be affected by the use of a specific hydro solver. Again, the results presented in Kereš et al. (2012) refer to the central galaxy and make no attempt to subtract out the contribution of poorly resolved haloes, so a straightforward comparison is not possible. However, we computed at $z=0$ the mass of central cooled gas concentrations (using the ‘star’ particles within $1/10$ of the virial radius) and compared it with the galaxy mass/halo mass relation of (Kereš et al. 2012, their fig. 4), finding within the statistics good agreement with the SPH result of that paper. We also checked density and temperature profiles of haloes, and found them to be roughly consistent with that presented in figs 8 and 9 of that paper, and inconsistent with the drop in temperature found in *AREPO* haloes. As expected, despite the different ‘star formation’ algorithms used, and in absence of effective thermal feedback, we obtain very similar results when we use the same code.

Intriguingly, both *AREPO* and SAMs predict, with respect to SPH, higher cooled masses and cooling rates at low redshift/large halo mass. We find that SAMs overpredict cooling rates by a factor of ~ 2 and cooled masses by a much smaller factor, while Kereš et al. (2012) report much larger differences between the two codes. However, for a proper comparison, one should check what is the difference when merging is neglected and the contribution from poorly resolved haloes is subtracted out. Regardless, a conclusive assessment of the accuracy of different implementations of hydrodynamics is necessary before firm conclusions on the behaviour of cooling in massive haloes can be reached.

6 SUMMARY AND DISCUSSION

We have tested cooling models embedded in three widely used SAMs (De Lucia & Blaizot 2007; Monaco et al. 2007; Benson 2012) by comparing their predictions to N -body hydrodynamical (SPH) simulations of radiative cooling in cosmological DM haloes. With respect to previous papers that performed similar tests (Benson et al. 2001; Yoshida et al. 2002; Helly et al. 2003; Cattaneo et al. 2007; Viola et al. 2008; Saro et al. 2010; Lu et al. 2011; Hirschmann et al. 2012), our simulations are improved in several ways: (i) we used a mass and force resolution sufficient to fully resolve the cooling region in all haloes larger than $5 \times 10^{10} M_{\odot}$; (ii) we computed cooling rates and cooled masses subtracting out the contribution of cooling in poorly resolved haloes; (iii) by also running a simulation where cooling is switched on at $z \sim 2$, we were able to test cooling models exactly in the redshift range where they are expected to be valid, with no influence from high-redshift overcooling; (iv) using a suitable formulation of SPH and a ‘star formation’ algorithm (as in Viola et al. 2008; Saro et al. 2010) to treat cooled particles as collisionless, we were able to limit numerical cooling and speed up the simulation considerably. As a word of caution, we recall that the simulations that we compare to SAMs have been carried out using an SPH hydrodynamical solver. Significant differences have been reported for the amount of star formation predicted by a simulation

based on the *AREPO* code that uses a Eulerian scheme (Kereš et al. 2012; Nelson et al. 2013); so some of our results, especially those relative to massive haloes, may be affected by the specific hydro solver we use.

We confirm that, overall, cooling models are able to approximately predict the correct amount of cooled mass. When cooling is active since the start of the simulation (the COOL simulation), median values are recovered at $z=0$ to within 20 per cent and with a similar scatter. At higher redshift, median values agree to within 40 per cent. For individual haloes, the worst discrepancies are rarely larger than a factor of 2. However, when the cooled mass in the COOL simulation is limited to properly resolved haloes, cooling models in SAMs underestimate the amount of cooled mass in the infall-dominated regime by ~ 20 –40 per cent at $z=2$. This difference is much less visible when the contribution of poorly resolved haloes is included, as the numerical underestimate happens to compensate the theoretical one.

In some models (*delucia* and *morgana*, and *galacticus* when an isothermal gas profile is used), cooled masses for more massive haloes tend to be higher than those found in the simulations at $z=0$. While this difference is found in the same range of halo masses where Eulerian and Lagrangian codes give different results (Kereš et al. 2012), this confirms the findings of (Saro et al. 2010, see also Hirschmann et al. 2012), who compared predictions from the *delucia* model with cooling rates from a simulation with very similar setting as our COOL simulation. This is a very relevant point: late cooling in massive haloes deposits mass in the central galaxy, which is typically a bright elliptical galaxy. The stellar populations of these galaxies are observed to be very old (e.g. Thomas et al. 2005), so quenching this cooling is necessary for any successful galaxy formation model. AGN feedback in the so-called radio mode (Bower et al. 2006; Croton et al. 2006) is typically advocated to produce this quenching; however, the fraction of galaxies with detectable radio emission associated with the AGN activity necessary to perform this quenching is higher than what is observed and shows the wrong dependence as a function of halo mass (Fontanot et al. 2011). As noticed also in Paper I, a cooling model that produces too strong a cooling flow at late times would require stronger AGN feedback to maintain quenching. For the *delucia* model, Saro et al. (2010) showed that the higher cooling rates with respect to results from simulations are due to the assumption of isothermal gas density profile. The same trend is shown by *morgana*, which assumes a hydrostatic density profile with a shallow inner slope. Clearly, the different type of integration used in the *morgana* cooling model causes the same trend without assuming a singular profile. However, before reaching a firm conclusion on this point, it is necessary to fully understand the role of the hydrodynamic scheme.

Consistently with the results of Paper I, the *galacticus* model with a cored gas profile underpredicts cooling flows in massive haloes at late times with respect to simulations. This prediction depends significantly on the assumed gas density profile: using an isothermal gas density profile generates predictions that are very similar to the *delucia* model and in much better agreement with the simulation, while large core radii are strongly disfavoured.

Cooling rates found in the simulations are recovered in the models with larger scatter, about a factor of 2. Scatter is larger for the *morgana* model, and part of it is due to the assumption in *morgana* that cooling is quenched during major mergers. We find no such trend in the simulations, and this is likely due to the persistence of cooled condensations during mergers. When the assumption of quenching at major mergers is dropped, *morgana* predicts cooling rates with less scatter, but is still larger than the other models.

Comparing models and simulations when cooling is switched on at $z \sim 2$ probably represents the cleanest test of the cooling model. In this case, in both the simulation and models, the baryons associated with haloes are not pre-processed by runaway cooling, they are all located in the hot atmospheres when cooling is switched on. Also, the contribution of cooling from poorly resolved haloes is very small for several halo dynamical times, and the deposition of cold gas is much less affected by all the numerical issues connected with the infall-dominated regimes. In our COOLZ2 simulation, the cooled mass is found to accumulate rapidly. At $z = 1.5$, *morgana* is able to predict the right amount of cooled mass, though with substantial scatter; *delucia* is biased low by ~ 20 – 30 per cent; and *galacticus* is low by a factor of 2, even when an isothermal gas profile is assumed. This confirms the findings of Viola et al. (2008), obtained with static haloes and using gas profiles that closely matched the simulated ones just before cooling is switched on.

Although switching on cooling at a given redshift provides a clean test of cooling models, one can wonder whether such conditions occur in realistic cases, and therefore, whether the difference among models, which is much less evident in the COOL simulation, should be a cause for concern. When cooling is active since the beginning, most baryons have already cooled by $z = 2$ (especially in relatively small haloes). The effect of this overcooling is visible, for instance, in Fig. 6, where cooling rates in the example Haloes 0 and 1 hardly exceed the value of $\sim 200 M_{\odot} \text{ yr}^{-1}$. These deposition rates can translate at best into star formation rates of the same order, which would be typically lower than the several hundred $M_{\odot} \text{ yr}^{-1}$ measured for massive star-forming galaxies at the same redshift. In the COOLZ2 simulation, cooling rates are higher by a factor of ~ 5 when cooling is suddenly switched on. Feedback from massive stars and accreting black holes is responsible for limiting overcooling at high redshift, and this makes more gas available at lower redshift. Moreover, as commented in Section 4.2, halo mergers or, more likely, feedback from star formation and AGN will likely be responsible for episodic quenching of catastrophic cooling, so the setting of the COOLZ2 simulation may be a good approximation for cooling flows in massive haloes at the peak of cosmic star formation, after a quenching event.

This is the first time that several cooling models running on the same merger trees are compared with a cosmological hydrodynamical simulation. Despite the simplified setting used in these simulations, we believe that they provide an important benchmark test for cooling models embedded in SAMs. Indeed, while energetic feedback from stars and AGN, which almost certainly plays a crucial role in shaping the properties of galaxies, is the most important contributor to the variance among model predictions (Fontanot et al. 2013), an accurate calibration of the condensation of gas in the central galaxy is desirable to remove unwanted sources of inaccuracies. Merger trees and results from the simulations presented in this paper are available to interested modellers upon request.

ACKNOWLEDGEMENTS

We thank Giuseppe Murante for his help in running the simulations and for discussions. We thank Volker Springel, Richard Bower, and Michaela Hirschmann for their feedback on this paper. Simulations were run on SP6 at CINECA thanks to the ISCRA-B project ‘Radiative cooling in galaxy formation: a numerical benchmark test for galaxy formation models’. PM and SB acknowledge financial contributions from the European Commission’s FP7 Marie Curie Initial Training Network CosmoComp (PITN-GA-2009-238356), from PRIN MIUR 2010-2011 J91J12000450001 ‘The dark Universe and

the cosmic evolution of baryons: from current surveys to Euclid’, from PRIN-INAF 2009 ‘Towards an Italian Network for Computational Cosmology’, from ASI-INAF agreement I/023/12/0, from PRIN-MIUR09 ‘Tracing the growth of structures in the Universe’, and from FRA2009 and FRA2012 grants of the Trieste University. GDL acknowledges financial support from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 202781. FF acknowledges financial support from the Klaus Tschira Foundation and the Deutsche Forschungsgemeinschaft through Transregio 33, ‘The Dark Universe’.

REFERENCES

- Benson A. J., 2012, *New Astron.*, 17, 175
 Benson A. J., Bower R., 2011, *MNRAS*, 410, 2653
 Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, *MNRAS*, 320, 261
 Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, *ApJ*, 599, 38
 Benson A. J., Borgani S., De Lucia G., Boylan-Kolchin M., Monaco P., 2012, *MNRAS*, 419, 3590
 Binney J., 1977, *ApJ*, 215, 483
 Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *MNRAS*, 370, 645
 Cattaneo A. et al., 2007, *MNRAS*, 377, 63
 Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
 Croton D. J. et al., 2006, *MNRAS*, 365, 11
 Cui W., Borgani S., Dolag K., Murante G., Tornatore L., 2012, *MNRAS*, 423, 2279
 De Lucia G., Blaizot J., 2007, *MNRAS*, 375, 2
 De Lucia G., Boylan-Kolchin M., Benson A. J., Fontanot F., Monaco P., 2010, *MNRAS*, 406, 1533 (Paper I)
 Dekel A., Birnboim Y., 2006, *MNRAS*, 368, 2
 Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, *MNRAS*, 405, 2161
 Fakhouri O., Ma C.-P., 2008, *MNRAS*, 386, 577
 Fontanot F., Pasquali A., De Lucia G., van den Bosch F. C., Somerville R. S., Kang X., 2011, *MNRAS*, 413, 957
 Fontanot F., De Lucia G., Benson A. J., Monaco P., Boylan-Kolchin M., 2013, preprint ([arXiv:1301.4220](https://arxiv.org/abs/1301.4220))
 Haardt F., Madau P., 1996, *ApJ*, 461, 20
 Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., Pearce F. R., 2003, *MNRAS*, 338, 913
 Hirschmann M., Naab T., Somerville R. S., Burkert A., Oser L., 2012, *MNRAS*, 419, 3200
 Katz N., Weinberg D. H., Hernquist L., 1996, *ApJS*, 105, 19
 Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, 363, 2
 Kereš D., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, *MNRAS*, 425, 2027
 Lu Y., Kereš D., Katz N., Mo H. J., Fardal M., Weinberg M. D., 2011, *MNRAS*, 416, 660
 Monaco P., Fontanot F., Taffoni G., 2007, *MNRAS*, 375, 1189
 Monaco P., Sefusatti E., Borgani S., Crocce M., Fosalba P., Sheth R. K., Theuns T., 2013, *MNRAS*, 433, 2389
 Murante G., Calabrese M., De Lucia G., Monaco P., Borgani S., Dolag K., 2012, *ApJ*, 749, L34
 Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
 Nelson D., Vogelsberger M., Genel S., Sijacki D., Kereš D., Springel V., Hernquist L., 2013, *MNRAS*, 429, 3353
 Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, *ApJ*, 725, 2312
 Planck Collaboration, 2013, preprint ([arXiv:1303.5076](https://arxiv.org/abs/1303.5076))
 Reed D. S., Smith R. E., Potter D., Schneider A., Stadel J., Moore B., 2013, *MNRAS*, 431, 1866
 Rees M. J., Ostriker J. P., 1977, *MNRAS*, 179, 541

- Saro A., De Lucia G., Borgani S., Dolag K., 2010, *MNRAS*, 406, 729
 Scannapieco C. et al., 2012, *MNRAS*, 423, 1726
 Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, *MNRAS*, 391, 481
 Springel V., 2005, *MNRAS*, 364, 1105
 Springel V., 2010, *MNRAS*, 401, 791
 Springel V., Yoshida N., White S. D. M., 2001, *New Astron.*, 6, 79
 Springel V. et al., 2005, *Nature*, 435, 629
 Srisawat C. et al., 2013, *MNRAS*, 436, 150
 Stanek R., Rudd D., Evrard A. E., 2009, *MNRAS*, 394, L11
 Sutherland R. S., Dopita M. A., 1993, *ApJS*, 88, 253
 Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, *ApJ*, 621, 673
 Tornatore L., Borgani S., Springel V., Matteucci F., Menci N., Murante G., 2003, *MNRAS*, 342, 1025
 Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, *A&A*, 506, 647
 Viola M., Monaco P., Borgani S., Murante G., Tornatore L., 2008, *MNRAS*, 383, 777
 White S. D. M., Frenk C. S., 1991, *ApJ*, 379, 52
 Yoshida N., Stoehr F., Springel V., White S. D. M., 2002, *MNRAS*, 335, 762

This paper has been typeset from a \LaTeX file prepared by the author.