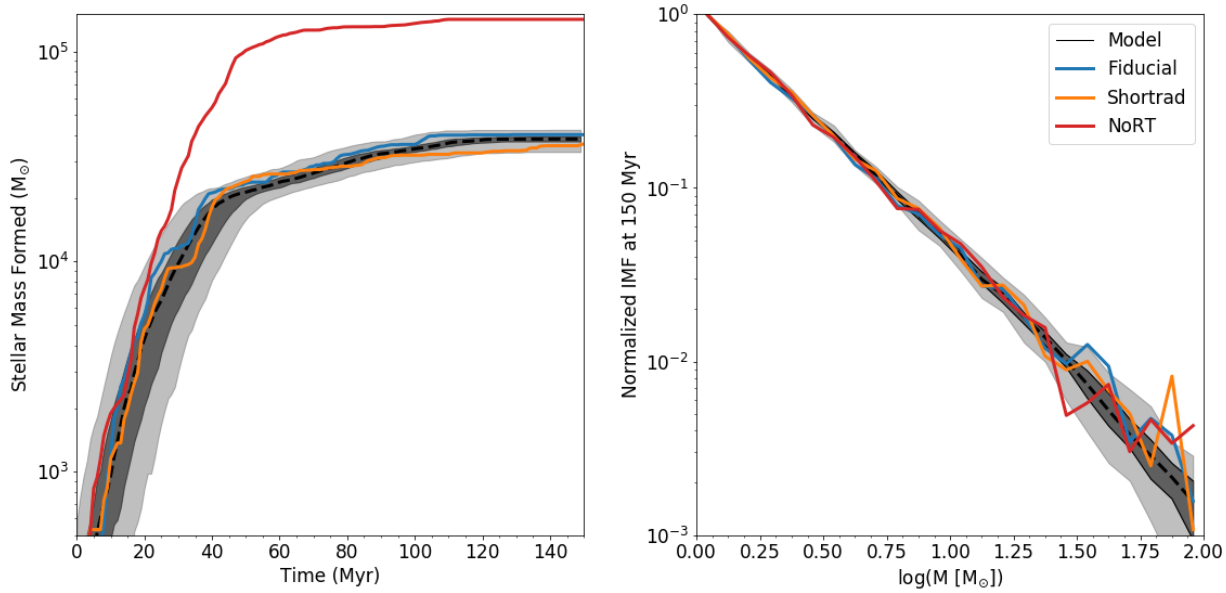
Response to Referee:

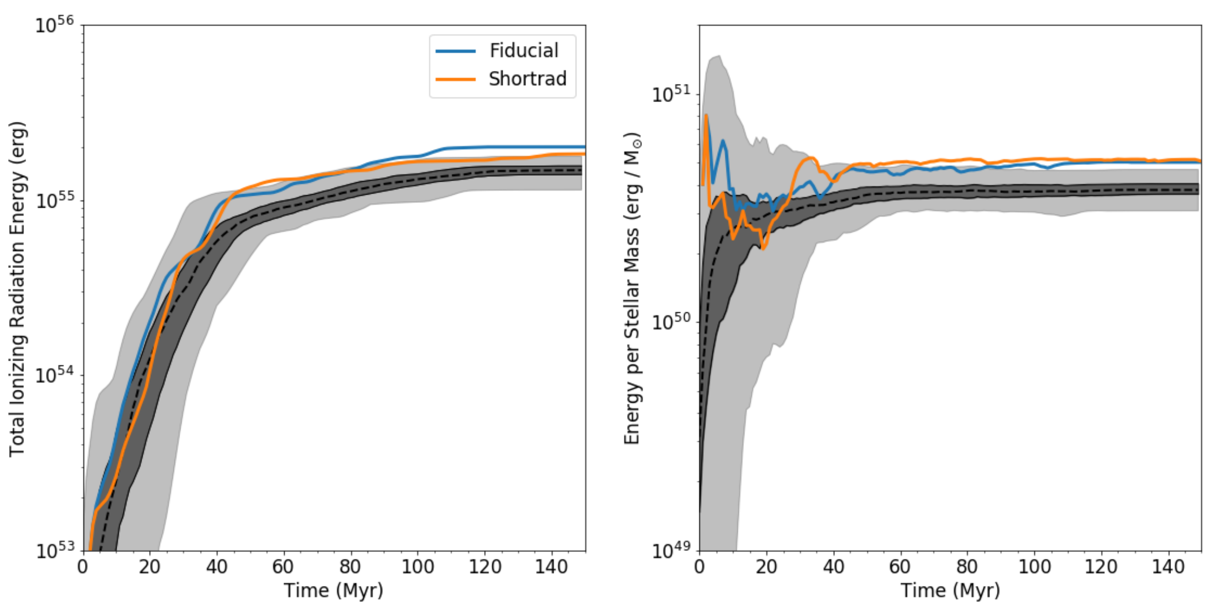
Response to Major Points:

Major Uncertainty #1: Stochasticity
Galaxy simulations are notoriously subject to stochastic effects (see, e.g. discussion in Keller+2018 arXiv:1803.05445), and this could be especially important in this particular point in parameter space. As a low-mass galaxy, the galaxy simulated has fairly low integrated star formation efficiency, and practically all star formation happens in the first dynamical time. I believe that stochastic effects are most pronounced for low SFE systems due to the greater proportional impact of individual star formation events, and it is easy to imagine that a similar effect could be at play here. The IMF sampling scheme used in the present study, while more realistic than past simulations, also introduces even more stochasticity. We are thus faced with the burden of demonstrating that the systematic differences between the different simulations are actually due to the different physics models used, rather than, for example, one simulation luckily drawing more than its share of 100msun stars and launching a massive wind. I am fairly convinced that this is not a problem for the "No RT" model, as the difference is so great and is well-documented in other studies, but it is much less clear for the fiducial vs. "Shortrad" comparison. The best way to address this is simply to run more simulations with different random seeds, as this would give some sense of the variance from both dynamical chaos and IMF sampling, however I understand that this may be onerous because these simulations are probably expensive. Failing this, the stochasticity from the IMF sampling can at least be evaluated to some extent by simply checking that the galactic ionizing flux-to-stellar mass ratios are comparable across the different runs, especially in the first few 10Myr during which most stars seem to form.

> Thank you for raising this point. The role of stochasticity in the evolution of low mass dwarf galaxies is an interesting and important consideration, and certainly a worry when interpreting the results of a small number of simulations. Unfortunately, running enough full simulations to constrain the effects of stochasticity on our results is computationally infeasible.
>
> However, we can understand these effects using a one-zone model matched to our simulations. We have separately developed a basic one-zone model that operates on a star-by-star basis with a stellar model having identical radiation properties to that employed in our hydrodynamics simulations. In the figures below, we compare the properties of our hydrodynamics simulations to the median properties of 100 matched runs of the one-zone model. In each model, we use a fixed input star formation history taken from our fiducial simulation. While this allows for a close comparison to the variance of the hydrodynamics simulations resulting from stochastic sampling, it likely underestimates the total variance since the SFR is held fixed, and not allowed to vary as a result of the sampling differences (e.g., the SFR of the one-zone model does not change if 100 one solar mass stars form instead of one 100 solar mass star).
>
> First, we compare the IMF properties across the simulations at 150 Myr:



The left panel shows the cumulative stellar mass formed in each simulation, while the right panel shows the normalized IMF at 150 Myr. The black line gives the median of the onezone models, with the dark shaded region showing the interquartile range, and the light shaded region the minimum and maximum of the one-zone models. As shown, the IMFs are well-converged across runs except perhaps at the highest masses, where there are still fluctuations.



More relevant to this report is the above figure. The left panel shows the integrated ionizing radiation energy emitted by all stars whose radiation is tracked (M > 8 M_sun), while the right panel shows the total energy per unit stellar mass formed. As the referee suggested, there is some variance here in the initial phase of star formation (~20-30 Myr). In particular, the fiducial model does have more radiation (in both panels) than the shortrad model for the first 30 Myr of evolution. However, these differences are small, and decrease over time in the runs. By 150 Myr, the integrated energy input per stellar mass for the two runs is nearly identical. While these differences likely do produce some differences in outflow properties between the two runs early on, they should not significantly impact the long term evolution.

This point has been added to the first footnote.

Major Uncertainty #2: Unresolved photon absorption
The simulations effectively follow a maximum density of 200cm^-3 due to the imposed star formation threshold, as evidenced by the phase structure plots shown in Paper I. As such, the internal structure of star-forming clouds is not necessarily resolved. The galactic ionizing radiation field could be sensitive to the characteristic density at which photons are absorbed, via the density-dependent recombination rate. One could imagine a pathological scenario where all photons are absorbed in ultra-compact HII regions close to massive stars where recombination is rapid, and none escape into the ISM at large. The extent to which such an effect is captured in the simulation would be a function of the maximum resolvable density in the simulation, which here is a function of resolution. Although I do not necessarily think that such a pathological scenario is likely to persist for any appreciable amount of time, based on the findings of high-resolution ray tracing studies of star-forming HII regions (e.g. Dale 2012-2017, Kim 2018), the fact remains that by introducing an effective density cutoff in the simulations we are setting the maximum density at which photons are absorbed. Therefore, we are making implicit assumptions about what is happening in dense HII regions, which can in principle affect the overall escape fraction of star-forming clouds. This is unfortunately a problem faced by all galaxy simulations that do not resolve the internal structure of dense clouds. For the purpose of evaluating the relative importance of "long-range" photons, I believe that it is important to determine the magnitude of this effect.

The most obvious way to do this would be through a resolution study: also try comparing fiducial and "Shortrad" runs at the 3.6pc and 7.2pc resolutions in the Paper I resolution study. Hopefully this wouldn't be too onerous, as I imagine the lower resolutions are significantly cheaper to run. Determine whether the differential effect of using the "Shortrad" approximation grows, decreases, or it is constant with resolution. If the differential effect is greater at lower resolution, that would suggest that the effects of unresolved HII region dynamics may be important. If that is the case, then the results of the study can only be interpreted as an upper bound on the relative importance of "long-range" ionizing photons.

Depending on the machinery available to the authors, the next best thing might be to simply post-process the resolution study runs in the Paper I appendix to determine the the characteristic density of UV photon absorption. One could imagine a histogram of the density at which photons are absorbed: if it does not shift appreciably with resolution, then it can be considered converged.

In the Milky Way, newly formed O stars have been observed to spend no more than 10 - 20 % of their main sequence lifetimes embedded in ultracompact HII regions within dense molecular clouds (Wood and Churchwell 1989). This likely represents an upper limit for the low metallicity regime we study here due to a lower dust content and increased stellar ionizing luminosity at fixed stellar mass (Geen et. al. 2015 includes an examination of how metallicity changes the impact of stellar radiation feedback in uniform media). Even these time scales are an order of magnitude longer than the free-expansion timescale for a Strömgren region, a delay that can be attributed to the continuing accretion flows into massive star forming regions (Peters et al. 2010). Therefore, we do not consider our neglect of this phase to be likely to lead to substantial errors at the galactic scale. We have added a brief discussion of this point to the text.

Unfortunately we cannot accurately determine the densities within which the photons are absorbed in post-processing, and this information is not recorded during the simulation.

Major Uncertainty #3: Fairness of comparison with short-ranged prescriptions
How comparable is the Shortrad setup to what is actually done in galaxy simulations that use the local Stromgren approximation? ie. is the 20pc cutoff actually representative? The Stromgren radius could potentially get quite large in a galaxy that is so dominated by diffuse warm neutral gas, so what if Shortrad is just ionizing significantly less material than the fiducial run due to having too small a cutoff? That is one possible explanation for the results obtained. If possible, please check the fraction of photons that are deleted at the 20pc radius. If the fraction is significant, then we aren't really making a fair comparison, because many of these local prescriptions expand their search for un-ionized material until all available photons have ionized something.

It is true that the cutoff of 20 pc is not necessarily representative of works that use a Strömgren-like approximation while allowing ionization to arbitrary distances (for example, as is done in the FIRE and FIRE-2 models). These works may avoid the discrepancy we identify in our submitted letter because of this. However, there are certainly many works that do not allow for this unbounded ionization, and only account for radiation feedback locally (e.g. Forbes+2016; Goldbaum+2016; Hu+2016,2017). Though we do recognize the definition of “locally” may differ between works, from a single cell / resolution element to an adopted maximum ionization radius (either a physical radius or a fixed number of cells / particles). It is the accuracy of this type of model that we focus on when comparing the shortrad and fiducial simulations. We argue this is a fair comparison.

Even more, neither the unbounded nor the localized Stromgren-like approximations account for angular effects. This makes their ionization severely mass-biased in non-uniform media. For example, photons from a star with a dense cloud to its left and diffuse gas to its right will not escape outward until the dense cloud is completely ionized. There is no opportunity for photons to escape asymmetrically through narrow channels in the ISM as these HII regions evolve spherically. This would also tend to over-ionize dense clouds far from stars that should only receive a fraction of their 4*pi solid angle. Hu+ 2017 discusses these points in greater detail. Hu+ 2017 uses a maximum cutoff of 50 pc in their Strömgren method to avoid mass-biased over ionization far from radiating stars.

We unfortunately do not record any information about absorbed or artificially deleted photons in our data outputs, and cannot estimate the fraction from a given source that are deleted at the 20 pc threshold.

We discuss this point in more detail in discussion of the paper.

Lastly, just a couple minor corrections:
Figure 1 legend: "Shortard" -> "Shortrad"
Section 3, paragraph 2: "black and blue lines" -> "orange and blue lines"

Corrected.

Additional Changes:

1) Abstract now emphasizes these results apply to low-metallicity, low-mass dwarf galaxies, which generalization required for more massive galaxies

2) All plots have been updated with additional run time of the shortrad simulation (up to 500 Myr) and the NoRT run (just over 100 Myr). This does not change the text.