We are grateful to the referee for devoting their time to our manuscript. The valuable comments and suggestions will help us to improve the paper.

We will here respond to the main comments made:

1) We are very sorry for not citing all of the relevant literature on scale dependent error growth, in particular in turbulence, and how to measure it numerically. We will fix this in the revised version.

Concerning the calculation of the error growth rates, we were indeed using a different scheme than proposed in Aurell et al PRL 77 1262 (1996): We create a perturbed field close to the reference field, we iterate them both for some transient time to allow the perturbation to re-direct into the locally most unstable direction, and then track the increase of the perturbation in time. Then (Fig.2a) we average over the error growth rate at given times after initialization, and plot it versus the mean error magnitude at that time (called $E(t)$ in Fig.2a). The reason for doing so is that in real forecasts, it is standard to study the average error after some given time. In particular, when performing the analysis for the ECMWF ensemble forecasts, we are only evaluating error growth experiments which have been performed by others, so that we do only have access to the error growth rate after fixed times, and not at given error magnitudes. We fully agree that the way how the averages are obtained generally will have an influence on the numerical results. We have therefore now repeated the error growth analysis for our model system in the way of Aurell et al. Qualitatively, the results are the same. With our method, the power law behavior $\lambda(E) \square E^{-\sigma}$ is slightly better than for Aurell’s method, which seems to have the tendency to resolve the Lyapunov exponents of the different levels. This smoothing of the error growth rate most certainly comes from the fact that, as suspected by the referee, our method mixes the scales: In the average over many error growth experiments, at some given time $t$ after initialization, the actual errors $E(t)$ have different magnitudes. In addition, we observe a slightly lower error growth rate for the same value of error magnitude, which we understand also by this mixing of scales: If $\lambda(E) \square E^{-\sigma}$, then $\langle \lambda(E) \rangle \times E^{-\sigma}$, where the average is assumed to be done over the error magnitudes which are found after fixed time $t$.

So our conclusion is that although a scale dependent error growth per se should be
studied in the way of Boffetta et al, a comparison with real weather forecast data can only be done by studying the time evolution of errors. We will add a corresponding discussion in the revised version of the manuscript.

2) line 33-34: Our statements about the “true” largest Lyapunov exponent being infinite were imprecise. They refer to a model where the scale dependent Lyapunov exponent reads $\lambda(\epsilon) \propto \epsilon^{-|\sigma|}$ and hence $\lambda \propto \infty$ for $\epsilon \rightarrow 0$. Evidently, such a behavior would be an idealization (approximation) since in real systems like in turbulence one would expect to have some cut-off at some small length scales with a cross-over to some finite $\lambda_{\text{max}}$. A comment like this was made in the publication Brisch & Kantz to which we refer in the manuscript, but I agree that we have to include it here as well.

3) We are sorry for ignoring relevant literature on scale dependent error growth and will give adequate merits to that in the revised version.

4) line 225: We fully agree, we erroneously forgot a prefactor $1/ \ln \epsilon$. We are grateful for pointing out this mistake.

5) Yes, we are talking here about the scale dependent error growth on scales which are already in the nonlinear regime. Due to the initial transient after the initialization of the perturbation, we do not observe the classical exponential error growth on very small scales, but it should be there. We will explore whether by the choice of even smaller perturbations, we can reach this regime numerically.

6) Unfortunately, it is true that for the ECMWF model results, we can not clearly distinguish between our conjecture of scale dependent error growth and a previous suggestion of how to fit it and how to interpret it. We have to understand better whether this is a consequence of (in term of scale dependent error growth) too coarse scales of the ECMWF model or of too large initial errors in the error growth experiments of ECMWF.

Concerning the referee’s conclusion we agree that our study of scale dependent error growth is not exactly in the spirit of FSLE. Inspired by the referee’s comments, we calculated the FSLE for our model, and we are considering to include the results in the revised version. However, the results of our “time after perturbation” approach do not differ qualitatively, and this is the only approach which works for the ECMWF model without modifying ECMWF’s code and running this model ourselves, which is out of reach to us.