An Evaluation of Filtering and Effective Resolution in the WRF Mass and NMM dynamical cores

Bill Skamarock, NCAR/MMM
Mike Baldwin, NOAA/NSSL
29 September 2003; revised 11 November 2003

With the recent introduction of the NCEP NMM dynamical core into the WRF software framework, we (Bill Skamarock and Mike Baldwin) believe it is appropriate, and indeed essential, to carefully evaluate the behavior of the new NNM core and the existing Mass core. These evaluations allow us to better understand the strengths and weaknesses of the cores, and give us guidance in determining appropriate WRF-core applications. Equally important, the evaluations can provide important guidance for making decisions regarding the allocation of the limited resources available to WRF project. Scientifically, we are motivated by results, presented by Mike Baldwin at the AMS NWP conference in August 2002 (Baldwin and Wandishin, 2002), showing that the accumulated precipitation produced by the Eta and NMM suffers from a significant lack of variance at the mesoscale compared with observational analyses. Interestingly, the WRF-Mass core does not appear to have this problem. In the more detailed evaluation of filtering in the mass and NMM cores presented herein, we find that the NMM core, as configured currently for operations, is significantly damped in its dynamics compared to observations and compared to all other nonhydrostatic mesoscale models we have examined (WRF-mass, COAMPS, MM5, MC2), consistent with the precipitation results. In this paper, we present the results of this initial evaluation followed by the significant questions raised by these results concerning WRF needs and plans.

1. Precipitation Spectra

At the AMS 15th NWP conference in San Antonio, and in various talks and seminars, we have shown results from Eta, NMM and various resolutions of the WRF model for a precipitation event occurring on 4 June 2002 depicted in Figure 1 (2002, Baldwin and Wandishin, AMS 19th WAF/15th NWP conference preprints, pp 85-88, San Antonio, TX, 2002). As is obvious in the figure, the Eta and NMM produce considerably smoother accumulated precipitation forecasts than the mass-core WRF or the observations. Depicted in figure 2 is the spectra computed from the precipitation, and it confirms the qualitative impression one derives from the precipitation forecasts. It shows that the Eta and NMM spectral slopes start deviating from the observations at wavelengths of the order 200 km (that is, around 15-20 Δx) whereas the WRF forecast spectral slopes start deviating at wavelengths of a little greater than 100 km for the 22 km forecast and perhaps 30-40 km for the 10 km forecast (that is, around 5 Δx).

Given our understanding of the Eta and NMM numerics and model configurations, we speculated that the considerable damping found in the Eta and NMM precipitation spectra at the shorter wavelengths was the result of the explicit damping in the model dynamics. Also pointing to this conclusion was the change in dissipation in the Eta-12

on 5 December 2001 following its blow-up shortly after it became operational in late November 2001; the background horizontal diffusion was increased by a factor of 8/3 and small-scale variance effectively disappeared from the precipitation forecasts (see the precipitation forecast archives at http://www.nssl.noaa.gov/etakf/qpfplots/). In a subsequent email discussion, it was argued that the smoothing of the precipitation fields might be produced by the convective parameterization (Eta and NMM use BMJ whereas WRF uses KF). Tests by Jack Kain and Mike Baldwin using different convective parameterizations and diffusion schemes in the Eta model did not entirely settle the issue. While there are indications that using the BMJ convective parameterization as opposed to the KF parameterization can somewhat reduce the small-scale variance in the precipitation forecasts, we believe the tests show that the bulk of the smoothing arises from the Eta and NMM models' explicit dissipation.

A critical issue here is the dissipation in the model dynamics, and a direct measure of dissipation in the model's dynamics is its kinetic energy spectra. With the introduction of the NMM in the WRF framework, we have examined the kinetic energy spectra in order to clarify these issues.

2. Kinetic Energy Spectra

2.1 Example Spectra

In contrast to the precipitation analysis, we do not have wind observations that allow computation of the kinetic energy spectra on a case-by-case basis. However, theory and observations suggest that the kinetic energy spectra should have a k^{-3} behavior at small wavenumbers (k) (large wavelengths) and should transition to a shallower $k^{-5/3}$ in the mesoscale (L $\sim$ few hundred kilometers; see Gage and Nastrom, 1986). The observational data are shown in figure 3. Gage and Nastrom also find that the spectra are relatively insensitive to altitude, latitude band, season, and variable (zonal or meridional wind, potential temperature). While a complete understanding of the dynamics underlying the change in spectral character has yet to be reached, the observations show that it does exist.

Figure 4 shows the kinetic energy spectra averaged over three 24-48h forecast periods (27, 29 and 31 July 2003) for the operational Eta and central US NMM forecasts produced at NCEP and for the 22 km CONUS WRF-mass forecasts produced at NCAR (see the appendix for a description of the spectra computations). The Eta and NMM spectra show no indication of a transition to $k^{-5/3}$, rather they have slopes greater than -3 for wavelengths smaller than approximately 300 km (approx 20 - 30 Δx). While the WRF-mass spectra does not show a strong transition to $k^{-5/3}$, the slope does not significantly steepen until around approximately 125 km (approx 5 Δx).

Figures 5 and 6 help to put the kinetic energy spectra in perspective. Figure 5 shows spectra from WRF BAMEX forecasts on 22, 10 and 4 km grids. As can be seen, the higher resolution models (10 and 4 km) show a much clearer transition to the shallower $-5/3$ spectral slope, although even the 4 km WRF mass model forecast does not reach

the $-5/3$ spectral slope. Figure 6 depicts all of the previous spectra along with the Gage and Nastrom GASP observational data (note that the WRF 22 km CONUS simulations from both 27-31 July and 2-4 June are plotted in red - the 22 km spectra overlie each other supporting a comparison between the BAMEX and late July spectra). The WRF mass model provides a better fit to the higher wavenumbers (smaller wavelengths) as the resolution is increased, whereas the Eta and NMM spectra are very damped.

We also examined the spectra of the WRF mass and NMM cores as a function of pressure (or height). Figure 7 shows the spectra integrated between different pressure levels for the WRF NMM core. As can be seen the energy decays rapidly as a function of altitude (pressure). This is not observed in atmospheric spectra (see Koshyk and Hamilton, 2001), nor is it observed in the WRF mass core, for which results are plotted in Figure 8, or other mesoscale models (we have examined MM5 and COAMPS in this regard).

Finally, in response to speculation that the spectra might be spuriously forced by topography and truncation errors associated with the WRF-mass discretization on the sigma coordinate grid, we have computed spectra for a the 10 km BAMEX domain moved to a location over the north-central Pacific Ocean, where there is no terrain. Figure 9 depicts the time-averaged spectra for a 28 October 2003 forecast over the Pacific and, for comparison, the June 1 BAMEX forecast spectra over the central US. The spectra are quite similar, demonstrating that the -5/3 behavior is not forced by the spectral distribution of the terrain nor by truncation errors in the discrete terrain transformation.

2.2 What Spectra Should a Mesoscale Model Produce?

Figure 10 shows spectra from a WRF mass-core BAMEX 10 km forecast (2 June 2003) at five times - 0, 6, 12, 18 and 24 h, starting with an initialization from a 40 km Eta analysis. As expected, there is a sharp drop-off in kinetic energy at wavelengths less than approximately 300 km in the initial state at 0 h. The WRF-mass core rapidly spins up the mesoscale spectra over the next six hours; over the following 18 h there are only small changes to the mesoscale portion of the spectra.

There is limited theory to guide us with respect to spin-up of the spectra, particularly with respect to the mesoscale portion. Tung and Orlando (2003), and to a lesser Gage and Nastrom (1986) and others, argue that the $-5/3$ spectra is produced primarily from a forward energy cascade (from larger to smaller scales). Lilly (1983, 1989) argues that energy input from small scales (convection) and an upscale cascade of a small amount of this energy (a few percent) produces the shallow spectra. Neither of these arguments suggest that the spin-up time should be long; a forward cascade should take on the order of an eddy turnover time, and input from small scales should result in larger-scale balanced flow in equally short time periods.

A more important question to ask is *"what does one observe in this spin-up period in the mesoscale forecast?"* Primarily, we see the formation of convection and mountain waves, along with a tightening of fronts, jet streams and other features in the flow

at all levels. The WRF mass core, and other models (MM5, COAMPS, etc.) produce fine scale structure dynamically consistent with mesoscale phenomena as we understand them. As expected, the unbalanced parts of the initialization also produce what we would call noise (gravity and acoustic waves). These waves persist for several hours but ultimately leave the domain or are dissipated in other ways. We do not see that they lead, in general, to significant spurious circulations.

Given that mesoscale models produce mesoscale structure in a short period of time when none exists in the initial conditions, and that this structure is observed, is physically realistic, and often verifies (for example, the forecasts of convective storms in 4 km WRF BAMEX forecasts were surprisingly good), it is clear that the spin-up of the mesoscale spectra is not entirely due to a downscale energy cascade from large scales; there is likely significant input from both small and large scales. Also, what we see in the models and observe in the atmosphere is a significant input of energy from external or physical mechanisms outside of the dynamical equations (convective heating, terrain forcing) that is not explicitly accounted for in theories. There is also dynamical forcing such as frontal collapse (which can happen very quickly (hours to a day) and cyclogenesis (a day to several days) that have strong three-dimensional components unaccounted for in 2D turbulence theory.

Based on turbulence theory, our understanding of mesoscale dynamics, and our evaluation of model spin-up and behavior, we believe that mesoscale models should reproduce the observed -5/3 spectral slope in their kinetic energy spectrum. By extension, mesoscale models should reproduce mesoscale structures that are dynamically consistent with observed mesoscale structures.

3. Filtering in the WRF NMM and Mass cores

Disregarding the filtering that arises from vertical mixing in the PBL physics, there are primarily two filtering mechanisms active in the WRF mass core.

(i) Implicit damping in the 5th order (horizontal) and 3rd order (vertical) advection schemes.

(ii) An explicit horizontal damping term with an eddy viscosity proportional to the horizontal deformation (with no minimum value for the eddy viscosity). This filter is used in the 4 km BAMEX simulations but not the 10 or 22 km simulations presented here.

Similarly, there are primarily two filtering mechanisms in the WRF NMM core.

(iii) An explicit horizontal damping term with an eddy viscosity proportional to the horizontal deformation (with a significant minimum value for the eddy viscosity).

(iv) An explicit horizontal divergence damping term with a constant eddy viscosity.

We have carefully examined results from the WRF mass core, and we believe it is the least dissipative model compared to existing cloud and mesoscale models (e.g.,

Takemi and Rotunno, 2003). The first two dissipation mechanisms (the RK3 scheme and the upwinding) have generally been sufficient to remove energy cascading from the large scales. At cloud-scales and below (where the turbulence becomes 3D), we have necessarily had to augment this energy removal with a turbulence parameterization.

The WRF NMM core is damped at a significantly higher rate than the WRF mass core, evidence the spectra in Figures 2 and 4. This increased damping comes from the large minimum horizontal eddy viscosity used in (iii) (approximately 17,000 $m^2\,s^{-1}$), and the large horizontal divergence coefficient used in (iv) (approximately 74,000 $m^2\,s^{-1}$). To further illustrate this difference in filtering, we configured the WRF mass core to filter in a manner similar to the WRF NMM core. Specifically, we used 2nd order centered advection (both vertical and horizontal, to remove the upwind filtering from the standard mass-core WRF configuration), we turned off all other computational spatial filters, we use the lower-bound viscosity from NMM in a second order horizontal damping term, and we use horizontal divergence damping with the same coefficient used in NMM. The results are given in spectra shown in Figure 11. WRF configured as NMM is significantly overdamped. WRF NMM spectra is also plotted in Figure 11. As can be observed, using the minimum horizontal eddy viscosity in the horizontal filter appears to under-estimate the total damping. In addition, we examined the impact of using the BMJ convective parameterization versus the Kain-Fritsch parameterization; a shown in figure 11, the kinetic energy is significantly effected only at the highest wavenumbers.

With regards to the difference in damping in WRF-mass and WRF-NMM forecasts, it should be appreciated that a significant amount of this difference lies in the scale-selective nature of the filters employed in the WRF-mass and WRF-NMM configurations. For example, a second order filter and a sixth order horizontal filter can be configured to damp $2\Delta x$ waves at the same rate. The second order filter, however, will then damp the 4, 6 and $10\Delta x$ waves at 4, 16 and 110 times the rate of the sixth order filter, respectively. Hence the low order filter, used in the WRF-NMM and Eta, tend to remove significant energy at what should be well-resolved scales on the model grid in addition to the necessary removal of energy close to the gridscale. The sixth order filter used in WRF-mass concentrates its energy removal at the highest wavenumbers, and in this sense is much more scale select than the second order filter used in WRF-NMM and Eta. A more complete explanation of the scale selectivity of filters is given in Dale Durran's book (Numerical Methods for Wave Equations in Geophysical Fluid Dynamics, Springer, 1991) in section 2.4.3.

Finally, we have also examined the history of lower bounds for the horizontal eddy viscosities used in the NCEP NMM and Eta models. The 22 km Eta, that was running operationally up through late November 2001, used an eddy viscosity lower bound value equal to $\sim 9,000\,m^2\,s^{-1}$. When originally introduced, the 12 km Eta used a value of $\sim 3,400\,m^2\,s^{-1}$, and this value was increased after the blowup of the 5 December forecast to $\sim 9,000\,m^2\,s^{-1}$ (with a corresponding decrease in effective resolution). The 8 km NMM uses a second order viscosity of approximately 16,000 $m^2\,s^{-1}$. In all other mesoscale models of which we are aware, the physical viscosity used in horizontal filters typically scales with some positive power of the grid spacing. *Paradoxically, the NCEP*

mesoscale models have been operationally configured to use the same or higher physical viscosities as their gridsize has decreased, with the result being that the models' effective resolutions have remained the same or decreased as their gridsize has decreased.

4. Issues

4.1 What Constitutes a Mesoscale Model?

An atmospheric forecast model produces realizations of the future state of the atmosphere. With respect to model resolution, mesoscale models should reproduce mesoscale structures that are dynamically consistent with observed mesoscale structures, and should resolve those structures at the finest scales representable on given grid. By extension, mesoscale models should reproduce the observed -5/3 spectral slope in their kinetic energy spectrum; if they generally reproduce observed structures, they should (and will) roughly reproduce the kinetic energy spectra.

By this measure, the WRF NMM core, and the operational Eta, are not mesoscale models as they are currently configured for operational use; they do not reproduce either the observed mesoscale spectra, nor by extension do they generally reproduce observed mesoscale structure (e.g., figure 1).

Furthermore, an important measure of model efficiency is the cost of a model simulation relative to the resolving capabilities of the model.

By this measure the WRF-NMM and the Eta, with their heavy filtering of mesoscale structure as documented by spectra and subjective forecast evaluation (e.g. figure 1), are highly inefficient.

With this understanding, the critical question with respect to the operational configurations of WRF-NMM and Eta is:

What is the scientific justification for running high resolution forecasts using numerical filters that systematically remove the resolution gained by the refined grid? In other words, if low-resolution forecasts produce similar verification scores, minimize false alarms and are smooth as desired by some forecasters, why not just run the forecast model at lower resolution?

4.2 Verification

Verification is a major research area in the WRF program and more generally in mesoscale and cloudscale meteorology and NWP research. As is shown in the spectra of the initial fields in a WRF mass core BAMEX forecast (figure 10), there is no significant energy in the mesoscale region of the initial state. The lack of energy in the analyses at the mesoscale is the result of a lack of observations at mesoscale resolution, and the smooth error covariances used in most existing data-assimilation systems; mesoscale observations, even if used, would likely be filtered from the analysis. Given the lack of observations and mesoscale analysis problems, we have no way to consistently verify

mesoscale structure in our forecast. Often times, however, these forecasts can be valuable to forecasters. How to measure that value, and whether or not the increased value is worth the increased cost, is the open question for mesoscale and cloudscale NWP.

Mesoscale verification is also problematic given the predictability timescale of mesoscale structures (e.g., convection). These timescales (tens of minutes to hours) are much less than mesoscale NWP forecast periods (1-2 days). Hence even with mesoscale observations, traditional point-wise verification methods are not likely to show improvement in forecast skill (see Baldwin and Wandishin 2002 NWP conference paper). Given this lack of formal predictability, it can be expected that the more filtered (damped) forecasts will verify better using traditional verification techniques. This appears to be the experience at NCEP with the Eta and NMM, and while it may justify increasing the damping in the forecast model, it does not justify running higher resolution models, because lower resolution models will greatly decrease cost and produce very similar verification scores.

In summary, traditional verification techniques are not likely to show increased skill in mesoscale forecasts, thus the increased cost of mesoscale forecasts will likely not be justified by traditional verification techniques. Other verifications, both subjective and objective, need to be developed and used to evaluate model performance, including spectra.

4.3 Applications

4.3.1 Cloud Resolving NWP

The WRF mass core has been tested in cloud-resolving ($\Delta x \sim$ few kilometers) NWP applications (BAMEX, SPC/NSSL Spring Program, etc.) idealized convective simulations of supercells and squall lines, etc). Generally speaking, the model is able to reproduce the results of other cloud models in idealized cases (and at somewhat coarser resolutions) and it has produced realistic, and sometimes verifiable, convective structure in the real-time NWP applications in support of field programs.

The WRF NMM core has not undergone any significant testing at these scales. Furthermore, extension of the current operational configuration to convection resolving resolutions ($\Delta x \sim$ few kilometers) will not produce convection. We have performed tests with the WRF mass core, using damping similar to that in the current operational NMM, with idealized convection-resolving simulations - no convection appears in these tests (the high level of horizontal damping and horizontal divergence damping removes any incipient convection).

Given these results, we believe that the WRF NMM core needs rigorous testing to determine if it is an appropriate model for mesoscale and cloudscale applications. The large damping used in the current configuration precludes evaluation of its applicability because the WRF NMM operational forecasts have no mesoscale structure (and no convective-scale forecasts have been presented). The NMM core may produce good results with reduced damping, but extensive evaluation is needed before any realistic mesoscale and cloudscale applications are warranted.

4.3.2 Hurricane WRF

There is a strong intent to use the WRF model as a hurricane model, by GFDL (and NCEP), by the Navy, and in university and government laboratory research programs. Nesting strategies will be used to allow convective resolutions ($\Delta x \sim$ few kilometers) within the next several years in operational use, with the goal of removing the need for convective parameterizations.

Given these plans, the need for a robust WRF model capable of resolving convection is a high priority. Questions regarding the convective resolving capabilities of the WRF NMM need to be answered before WRF NMM can be considered for WRF hurricane applications.

4.3.3 Ensemble Forecasting

There are many scientific questions that remain unanswered concerning the appropriateness of the ensemble WRF operational plan announced at the WRF June 2003 workshop. Given the strongly damped mesoscale forecasts produced by the WRF NMM, even more questions are raised. *Highly damped models are known to have significantly smaller rates of error growth than expected from theory or produced by other, less damped models. Is it appropriate to populate half the ensemble members with highly damped dynamics? Is it appropriate to use highly damped dynamics in any atmospheric model ensemble?*

Appendix

The spectra are computed from the model output velocity fields. We are computing 1D spectra along horizontal grid-lines and along constant pressure surfaces using the model forecasts winds (we have found no significant difference between the pressure or sigma surface spectra). In this way we crudely mimic the aircraft flight-track measurements used in the Gage and Nastrom observational analysis. For the WRF model, the lines are west-east (corresponding the the longest model dimension) and start 15 grid-points in from the lateral boundaries. We have also used north-south crosses and found no significant difference in the results. For the Eta and NMM model we use lines that are oriented SW-NE and NW-SE in order to sample along the shortest grid distance on the E grid. We have also computed spectra using a west-east orientation with the Eta and NMM model (which has an effective gridlength of $\sqrt{2}\Delta x$) and found no significant differences in the results. In a typical computation (e.g., figure 4), the spectra are averaged over approximately 60,000 lines for Eta, 33,000 lines for WRF NMM, and 21,000 lines for the WRF mass core 22 km forecasts (total over the three days - 27, 29 and 31 July). The forecasts are sampled every three hours over the 24h diurnal cycle, and the forecast period is typically from 24 to 48 to avoid spinup issues. The Gage and Nastrom data are from aircraft and tend to be from the upper troposphere and lower stratosphere. The observations cover all seasons and are from both midlatitudes and tropics. Gage and Nastrom find some differences between seasons and latitudes, but they tend to be small. In both the observations and forecasts, the largest wavelengths are only par-

tially sampled - we are running regional as opposed to global forecasts, hence variations at the smallest wavenumbers are to be expected. The spectra are computed in the standard way for 1D non-period functions; the data are first de-trended. Finally, the forecast domains for the different spectra differ. Given that the spectra from like models track each other (NMM and Eta, different WRF domains and resolutions), the different domains do not appear to skew the results in any misleading manner. Also, all domains are generally centered over the central US.

References

Baldwin, M. and M. Wandishin, 2002: Determining the resolved spatial scales of Eta model precipitation forecasts. Preprints, AMS 15th Conference on Numerical Weather Prediction, 85-88.

Gage, K.S., and G. D. Nastrom, 1986: Theoretical interpretation of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft during GASP. *J. Atmos. Sci.*, **43**, 729-740.

Koshyk, J. N., and K. Hamilton, 2001: The horizontal kinetic energy spectrum and spectral budget simulated by a high-resolution troposphere-stratosphere-mesosphere GCM. *J. Atmos. Sci.*, **58**, 329-348.

Lilly, D. K., 1983: Stratified turbulence and mesoscale variability of the atmosphere. *J. Atmos. Sci.*, **40**, 749-761.

Lilly, D. K., 1989: Two-dimensional turbulence generated by energy sources at two scales. *J. Atmos. Sci.*, **46**, 2026-2030.

Takemi, T., Rotunno, R. 2003: The Effects of Subgrid Model Mixing and Numerical Filtering in Simulations of Mesoscale Cloud Systems. *Monthly Weather Review*, **131**, No. 9, 2085-2101.

Tung, K.K., and W. Orlando, 2003: The k^{-3} and $k^{-5/3}$ energy spectrum of atmospheric turbulence: Quasigeostrophic two-level model simulation. *J. Atmos. Sci.*, **60**, 824-835.

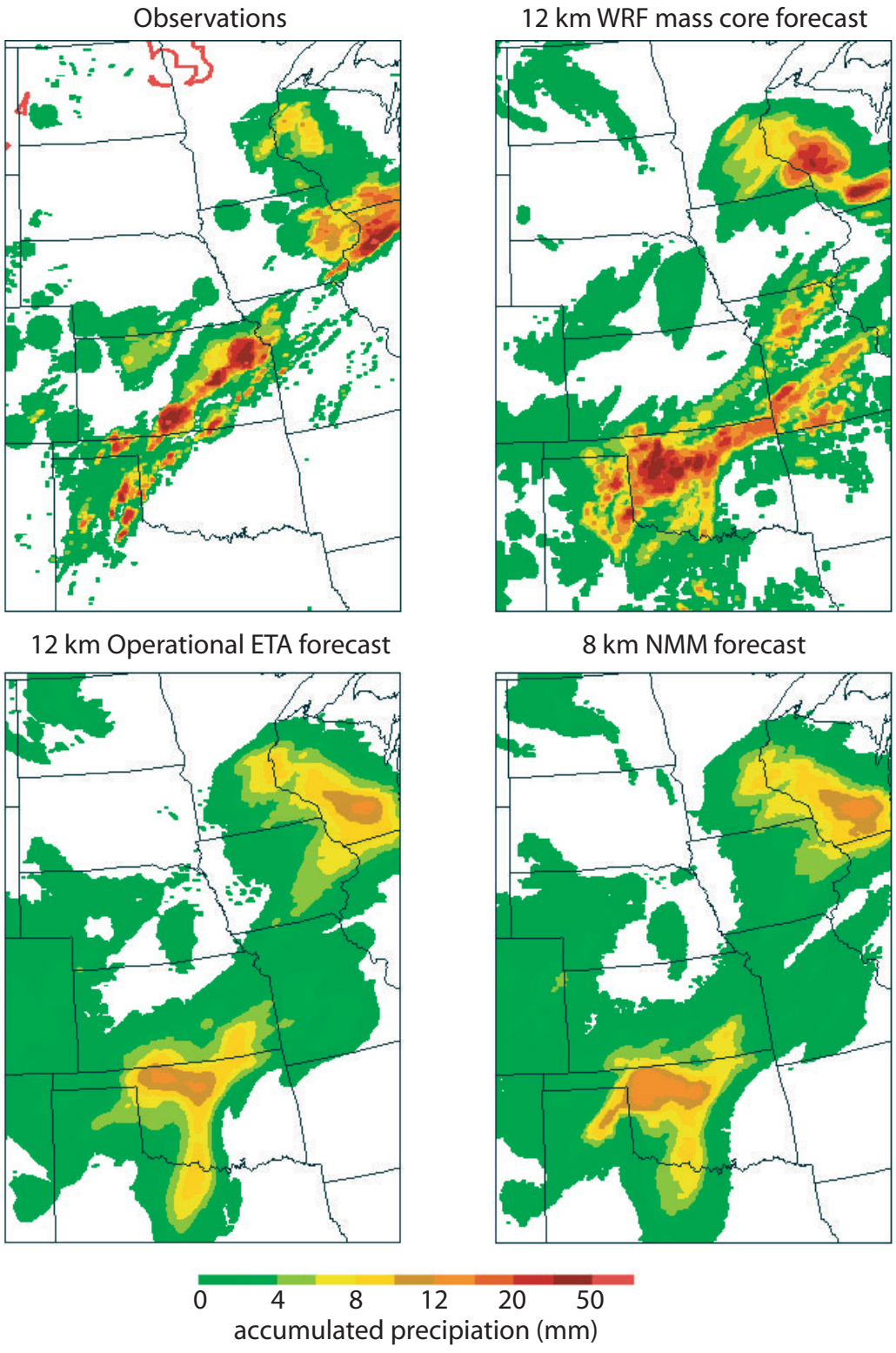

Figure 1: 3 hour accumulated precipitation valid 15-18 Z 4 June 2002.

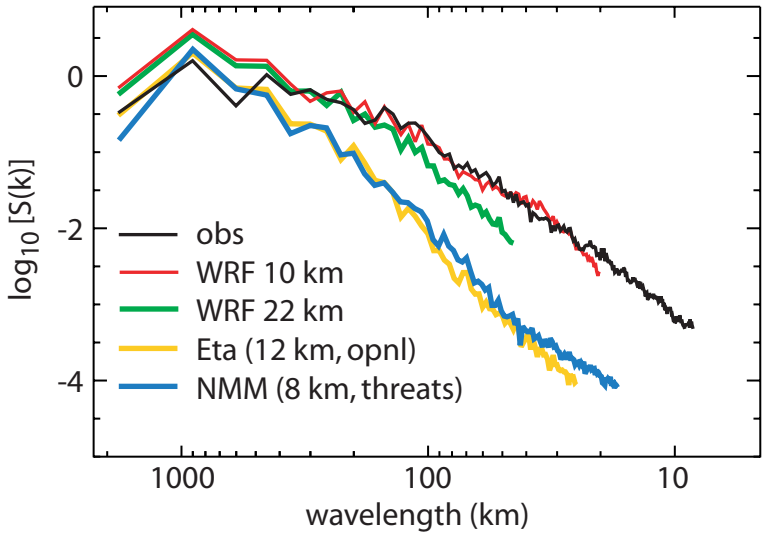

Figure 2: Spectra computed from the 3 hour accumulated precipitation given in figure 1 (15-18 Z 4 june 2002).

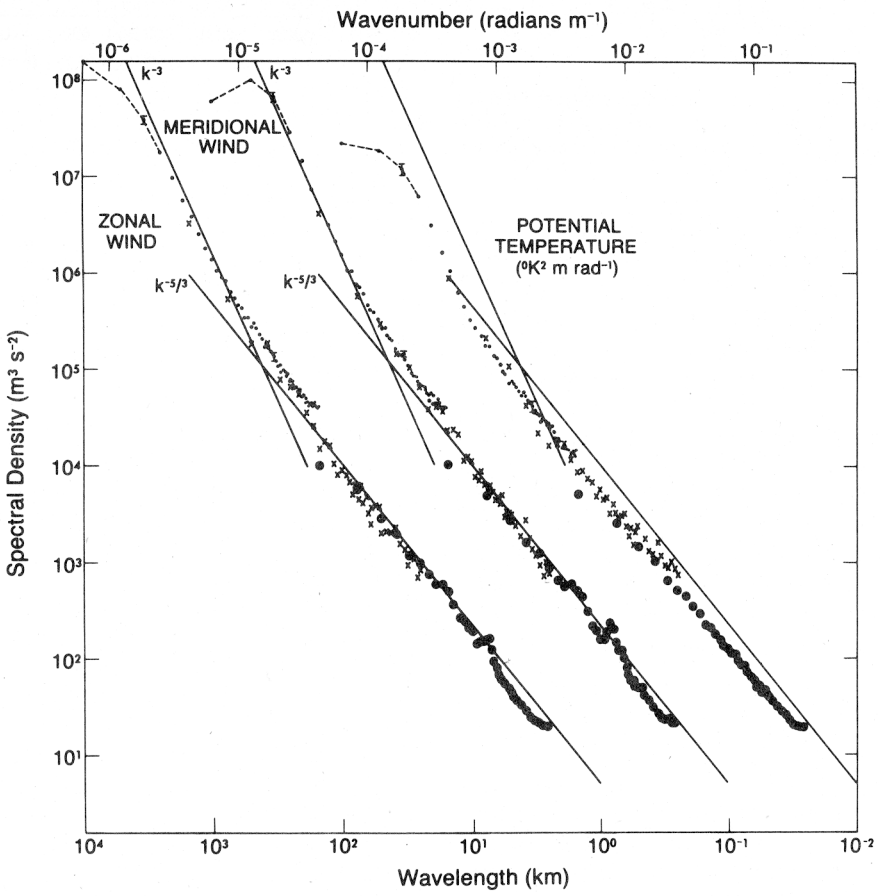

Figure 3: Spectra for the zonal and meridional velocity and potential temperature from Gage and Nastrom (1986). The spectra for meridional velocity and temperature are shift 1 and 2 decades to the right, respectively, for clarity.

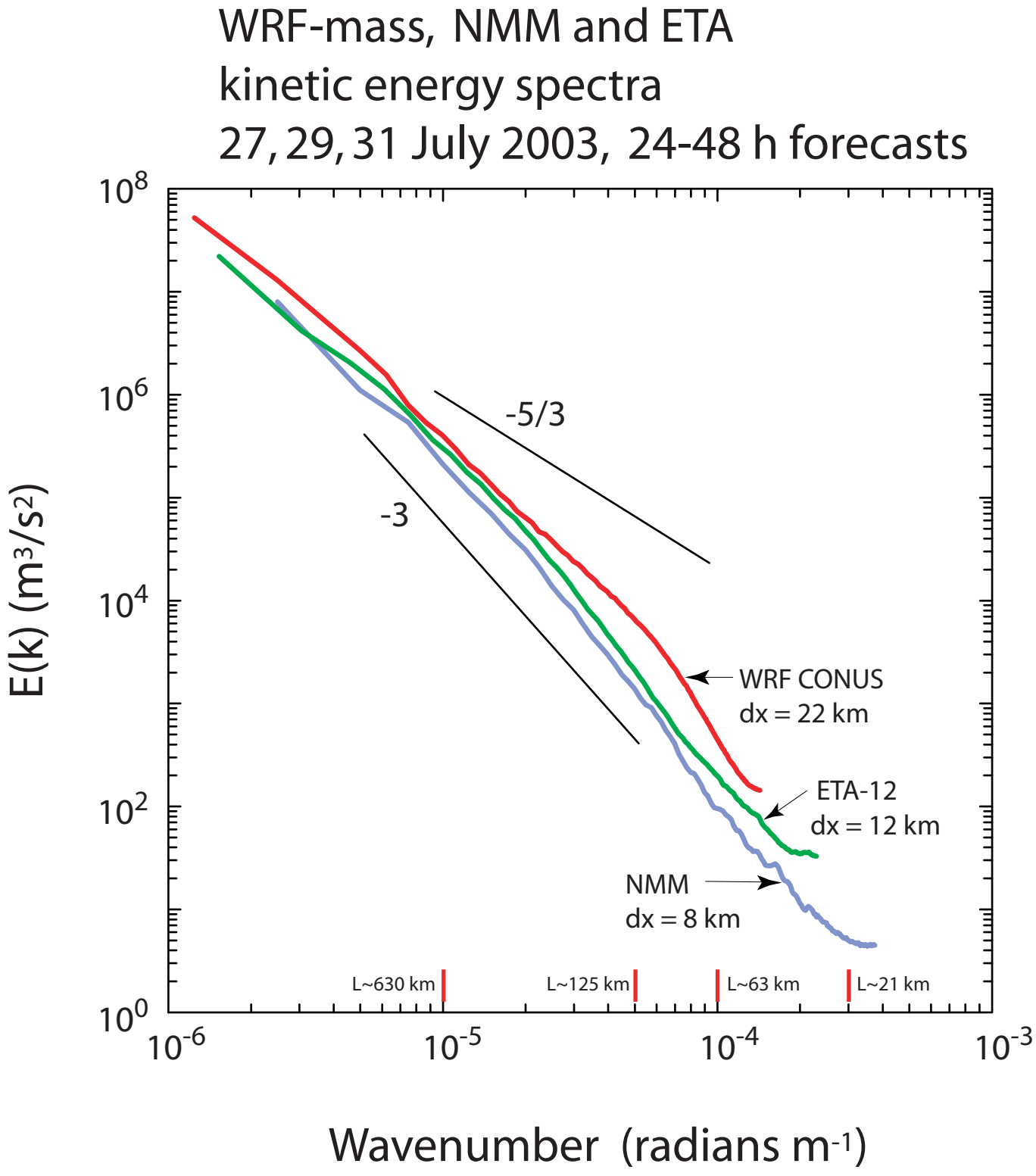

Figure 4: Kinetic energy spectra for ETA, NMM and WRF forecasts.

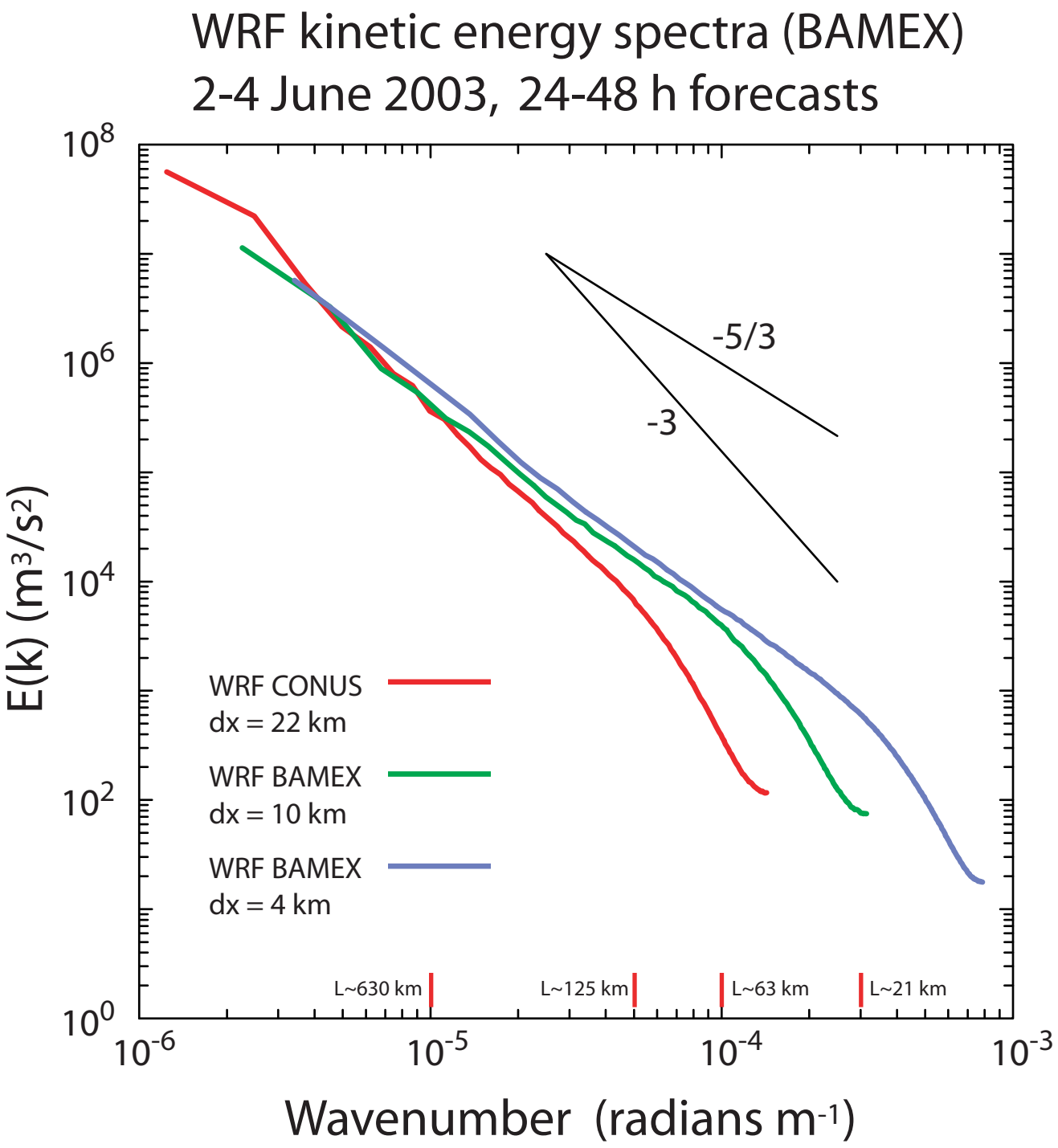

Figure 5: WRF mass core BAMEX forecasts in 22, 10 and 4 km grids.

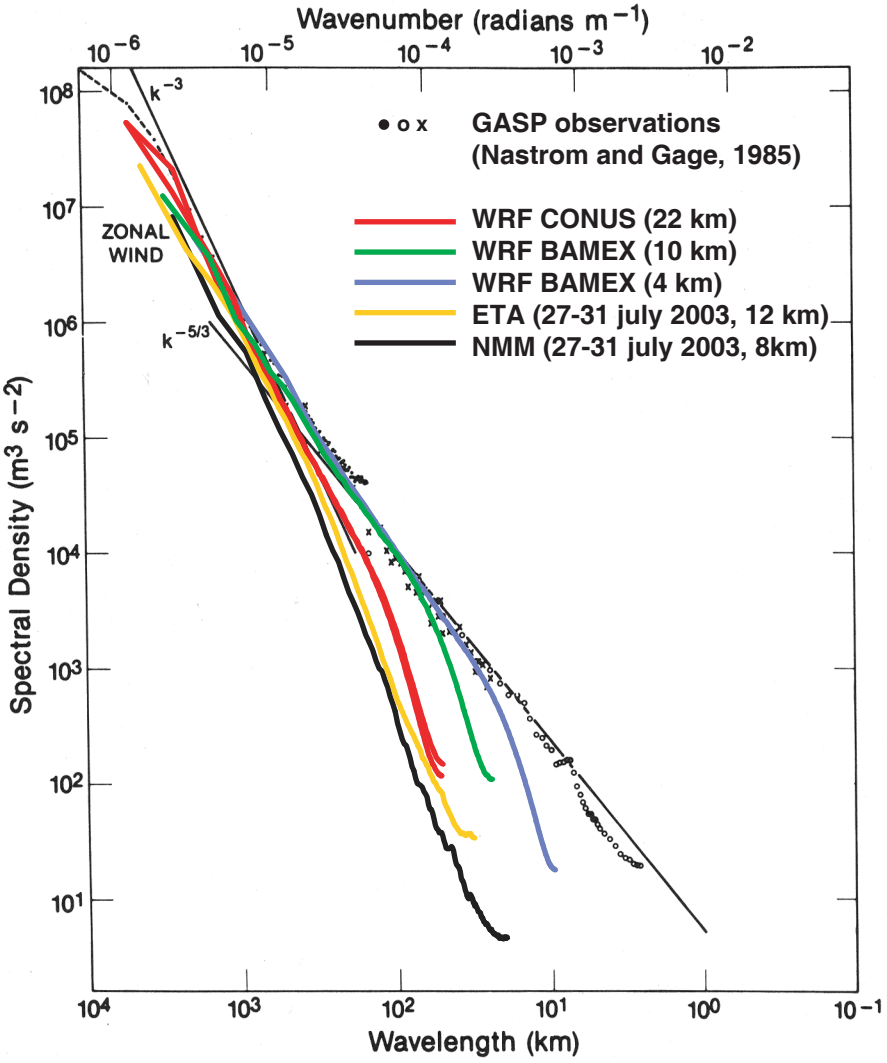

Figure 6: Forecast spectra plotted with the Nastrom and Gage (1985) observational spectra.

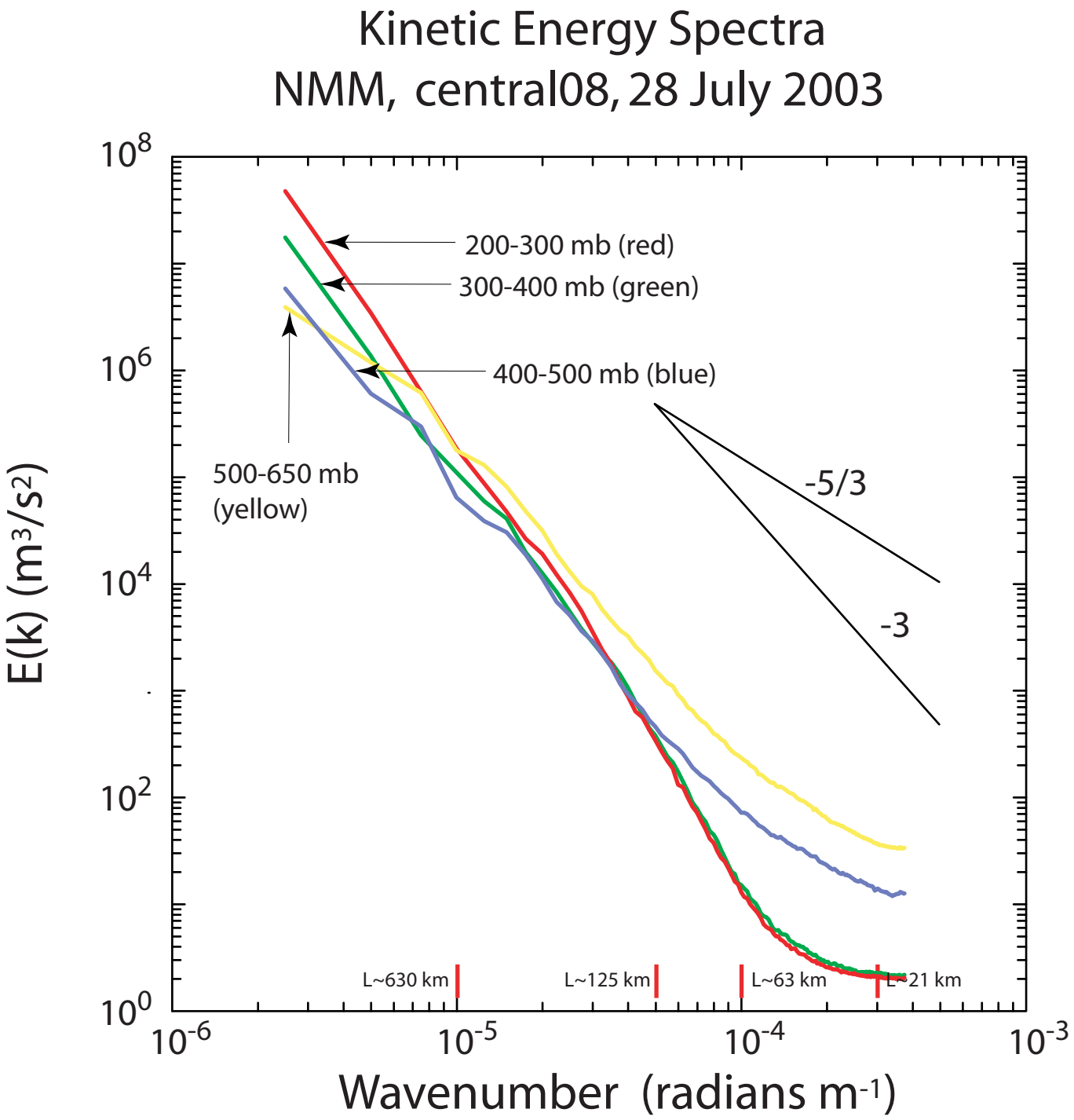

Figure 7: Kinetic energy spectra for NMM 24-48h forecast valid 28 July 2003. The spectra is integrated between pressure levels, and shows the significant decay in the mesoscale spectra with height.

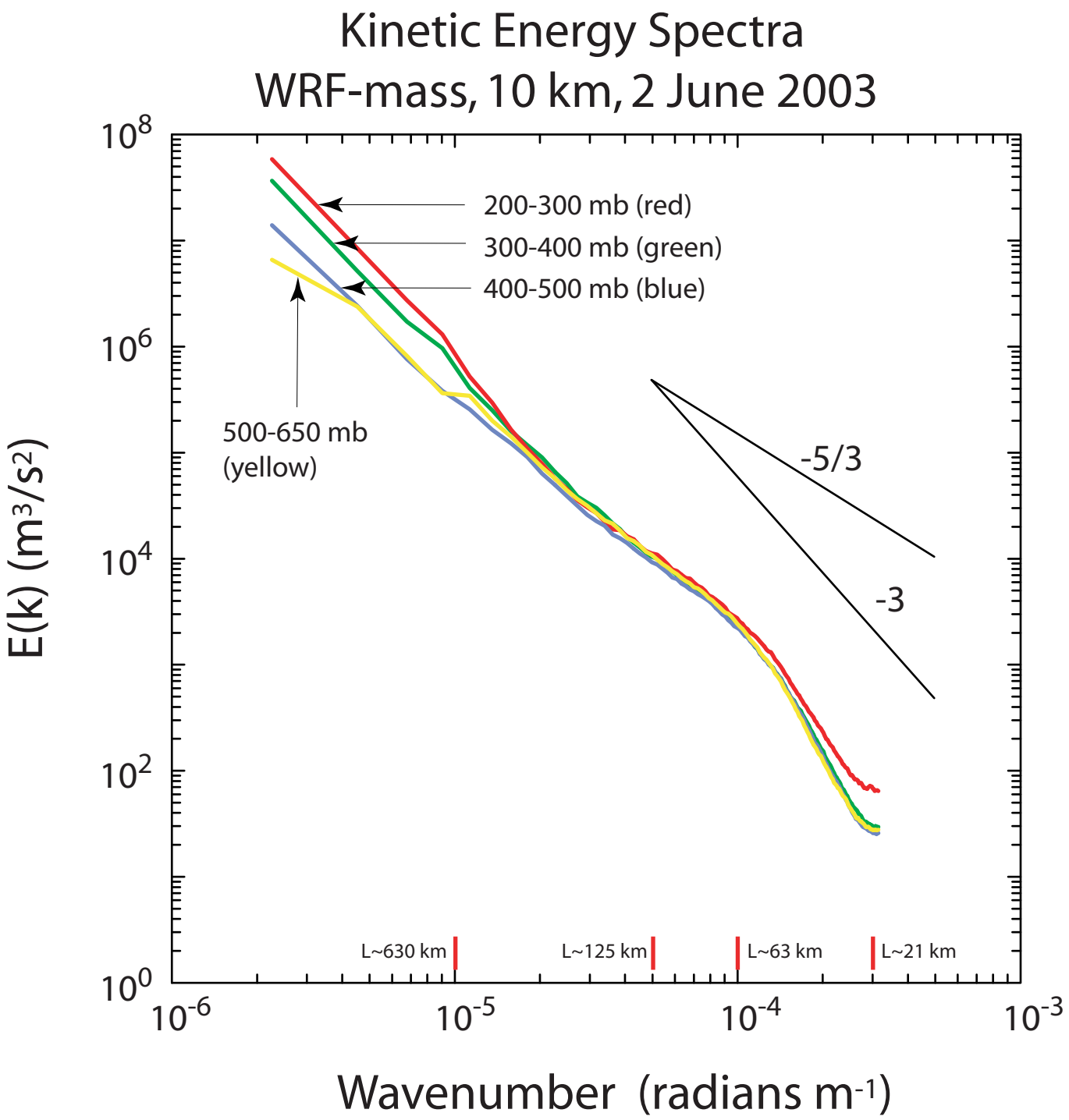

Figure 8: Kinetic energy spectra for WRF-mass 24-48h forecast started at 0Z 1 June 2003. The spectra is integrated between pressure levels, and shows that there is little difference in the spectra with height in the mesoscale.

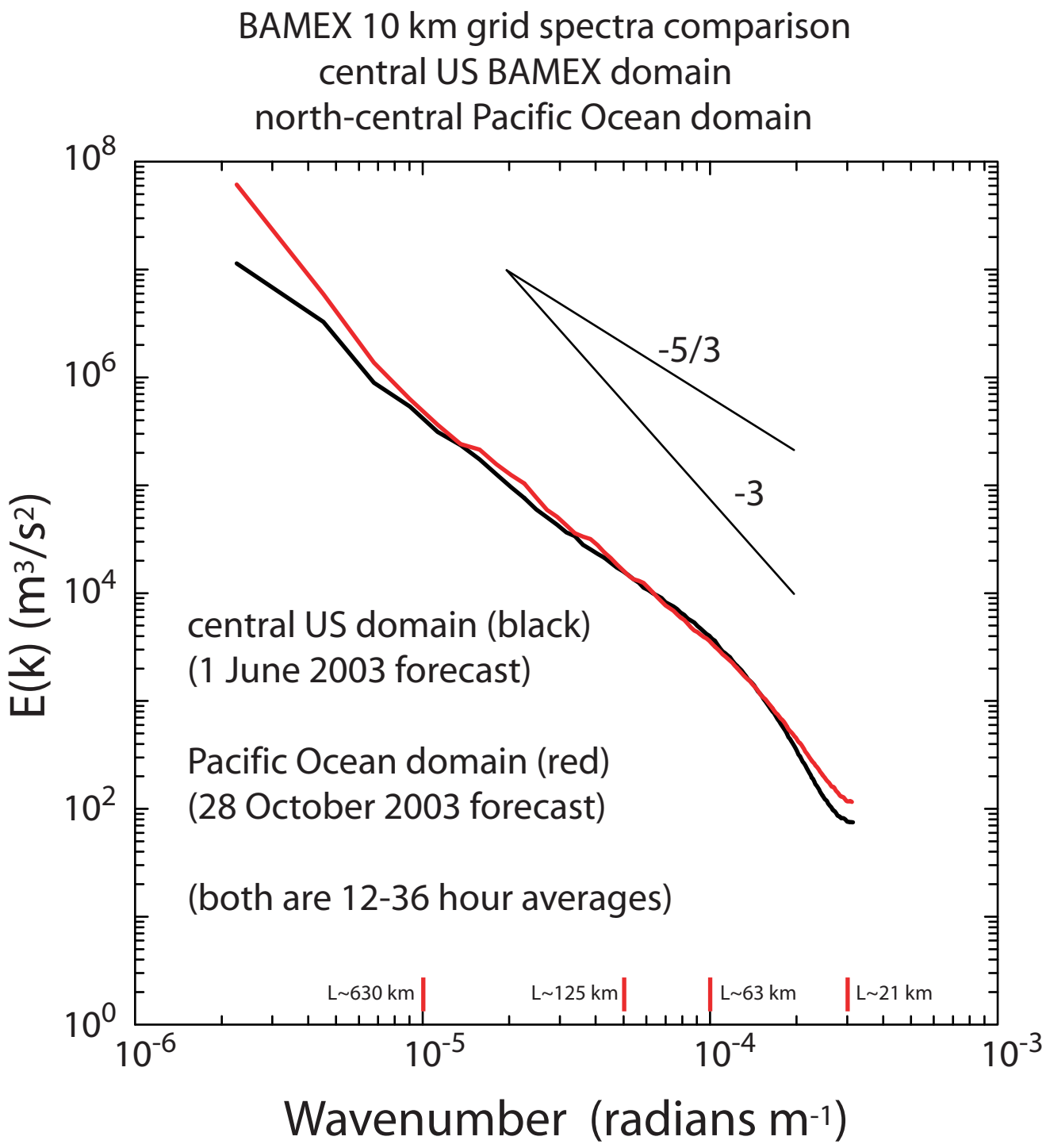

Figure 9: KE spectra for 28 October 2003 forecast using the 10 km BAMEX grid and configuration with the grid placed over the north-central Pacific Ocean.

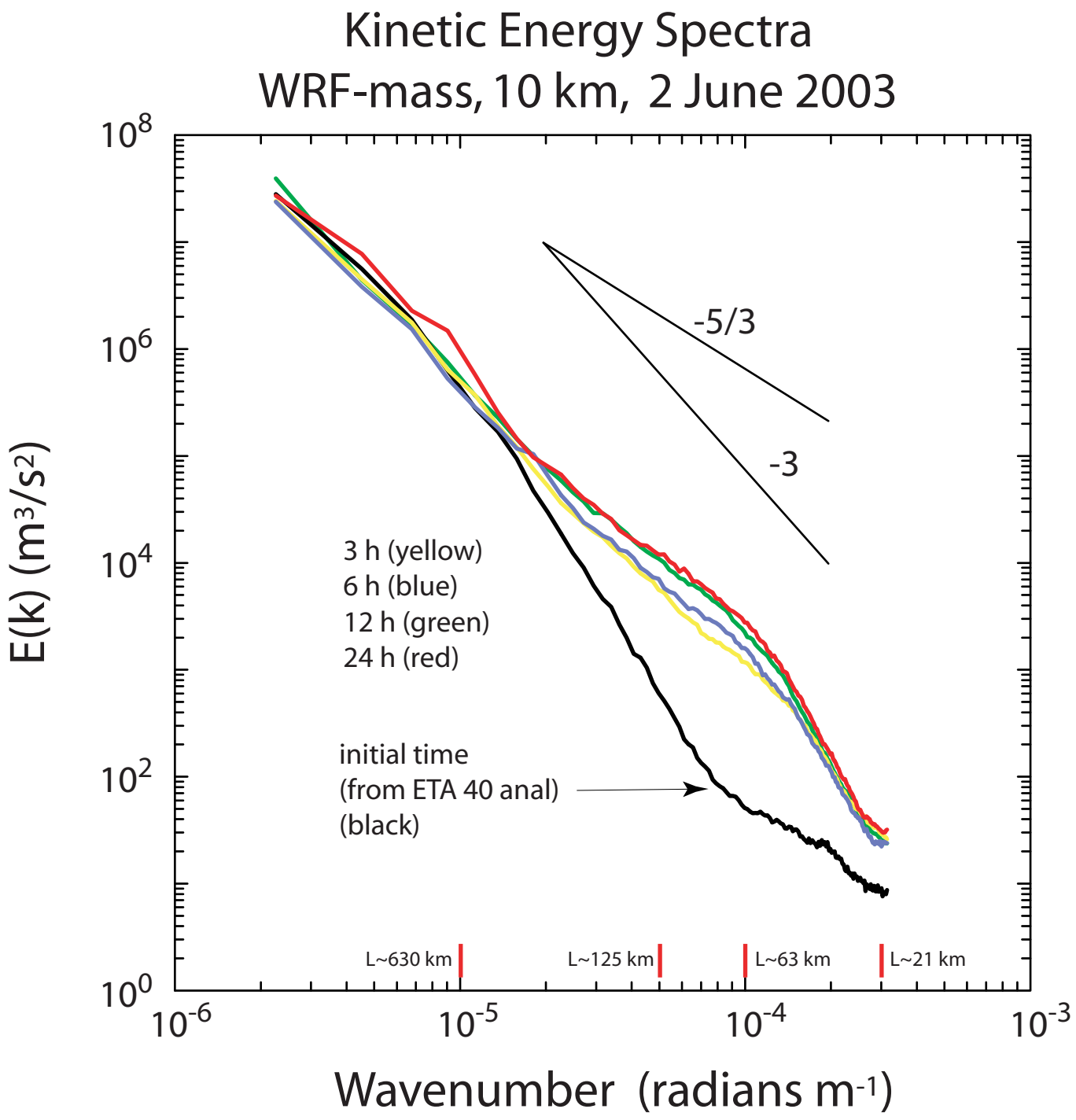

Figure 10: Kinetic energy spectra for WRF-mass 0-24 h forecast started at 0Z 1 June 2003, averaged over 200-500 mb (10 mb increments).

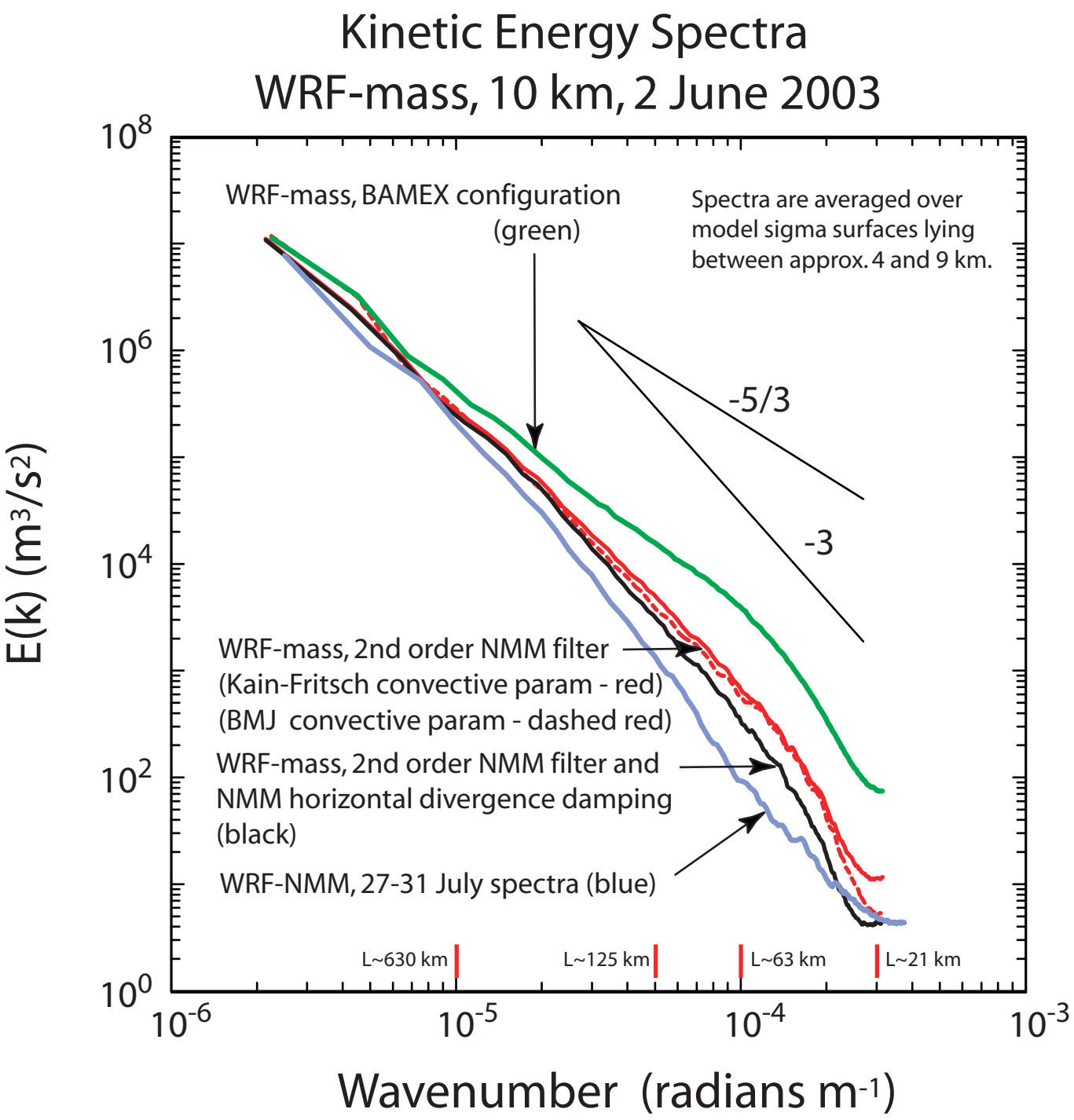

Figure 11: Kinetic energy spectra for WRF-mass 24-48h forecasts started at 0Z 1 June 2003. 2nd-order centered advection was used for NMM filter tests and all other spatial filters in WRF were removed. In the WRF-mass tests, the second order NMM eddy viscosity was fixed at the NMM lower bound of 17,000 m^2/s, and the eddy viscosity for horizontal divergence damping was 74,000 m^2/s. Using the lower bound for the eddy viscosity appears to underestimate the NMM damping (as suggested by the comparison with the NMM 27-31 July spectra). Additionally, the mesoscale is preferentially damped by the horizontal divergence damping (compare black and red spectra). Also, using the BMJ convective parameterization, as opposed to Kain-Fritsch, adds little to the damping except at the highest wavenumbers (compare the red solid and dashed spectra).