High Resolution Numerical Weather Forecasting to Aid AMOS

Kevin Roe
Maui High Performance Computing Center, Kihei HI

Mark Waterson
Boeing, Kihei, HI

ABSTRACT

The Hawaiian Islands contains a variety of microclimates in a very small region. Some islands have rainforests within a few miles of deserts; some have 10,000 feet summits only a few miles away from the coastline. Because of this, weather models must be run at a much finer resolution to accurately predict in these regions. NCAR's Mesoscale Model Version 5 (MM5) if run from a coarse 27 km resolution (surrounding an area of approximately 5000 by 5000 km) nested down to a 1 km resolution daily. Since the computational requirements are high to accomplish this in a reasonable time frame (as to still be a forecast) MM5 is run in parallel on MHPCC's IBM SP3. Utilizing 32 processors the MM5 model is run over the above conditions in approximately 7 hours.

These forecast have been in place over a year now and are being utilized by operators at the telescope operators on Haleakala, Maui. Forecast produced are for a 48-hour simulation; this forecast is available to operators by 10 AM and give forecasts out until the next day at 10 PM. This is enough time to give operators and managers time to reschedule their operations if unacceptable conditions are predicted. The products we currently provide are: temperature, wind speed & direction, clouds (at low, middle and high levels), relative humidity, rainfall, and a sounding (vertical profile of temperature, dewpoint temperature, and wind direction) for the Haleakala summit.

1. MOTIVATION

The telescope operations on Haleakala are highly dependent on weather conditions on the Hawaiian Island of Maui. If the wind speed is too high then the telescope cannot be utilized. Problems also exist if there are clouds overhead. Rainfall and relative humidity are also a factor in determining the capabilities of the telescopes. In order to effectively schedule telescope operations, an accurate weather prediction is extremely valuable. Current forecasts that are available from the National Weather Service (NWS) give good indications of approaching storm fronts but only at the coarse level (30-50 km resolution). Because of this and the location of the telescope on Maui this can be insufficient for their needs.

The additional benefit of the telescope operators having access to an accurate forecast (even for only a day in advance) is that they can still perform some scheduling. If a storm is predicted they can plan maintenance for this time period. This allows them to function more effectively by given them the capability to schedule downtime. This in turn saves them time, improves their operating efficiency, and in turn potentially save money.

2. NUMERICAL WEATHER MODELING

The numerical weather model (NWM) used for this project is NCAR's Mesoscale Model Version 5 (MM5) [1]. It was chosen because it has many desirable features:

(a) Multiple nested grid capability,
(b) The ability to include observational data into the model,
(c) The terrain portion of the model includes vegetation and land use data,
(d) The model is capable of being run in parallel on MHPCC'S IBM SP3,
(e) It is the mainstay NWM for the Air Force,
(f) It will grow into the Weather Research and Forecasting (WRF) Model [2].

The nested grid capability allows a coarse mesh to be run over a large area in little compute time while still being able to operate on finer resolution grids for specified locations of interest (i.e. the summit of Haleakala). The ability to include observational data allows the model to start with a better "first guess", which in turn allows the model to "spin up" quicker. Finally, the ability for the model to be run in parallel is important because it allows the

production of high-resolution output in a reasonable time frame (when the forecast is still a prediction). Benchmarks have been done for the parallel version of MM5 (using MPI), but it is not easy to make comparisons since the simulation's domains is significantly different than NCAR's choice of domains and resolutions. (See NCAR's web page on parallel performance at **http://www.mmm.ucar.edu/mm5/mpp.html**).

The best way to understand how the MM5 model operates is to explain the main routines it uses to accomplish a numerical simulation. Fig. 1 is a flow chart of the main routines used in the MM5 model. The TERRAIN program horizontally interpolates the regular latitude-longitude elevation and vegetation onto the chosen mesoscale domain; it then outputs data files that are used by the REGRID, NESTDOWN, and MM5 portions of the MM5 model. The REGRID program reads meteorological analyses on pressure levels and interpolates them from some native grid and map projection to the grid and map projection defined by TERRAIN; it then creates data files useable by RAWINS, LITTLE_R, and INTERPF. The RAWINS and LITTLE_R programs are designed to improve/enhance the first-guess meteorological data (usually received from REGRID) through the inclusion of observational data. The INTERPF routine handles data transformations that are necessary to put analysis data into a format useable by the mesoscale model. INTERPF ingests data from REGRID, RAWINS, or LITTLE_R, performs vertical interpolation, diagnostic computation, and data reformatting to create initial, lateral boundary, and lower boundary conditions for the mesoscale model. The MM5 program is the numerical weather prediction portion of the modeling system. The NESTDOWN program is to horizontally interpolate sigma-coordinate data from a coarse grid to a fine grid.
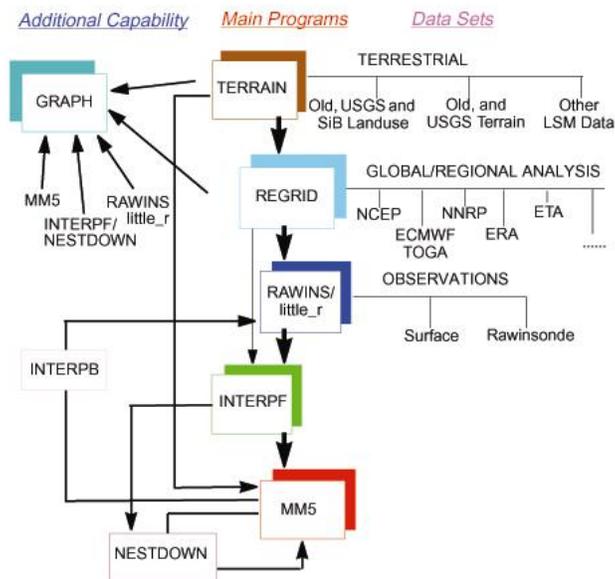

Fig. 1 MM5 Model Flow Chart

3. SETUP AND AREA OF INTEREST

MM5 is a nonhydrostatic, three-dimensional primitive equation model utilizing terrain-following sigma vertical coordinates [3]. In this simulation we will use:

1. 26 sigma levels from the surface to the 100 mb level with a bias towards levels below a sigma of 0.9 (close to the surface). High vertical resolution is needed at the lowest levels to resolve the katabatic flow and nocturnal inversion in the near surface layer [4,5].
2. Grell's cumulus parameterization for the 27 and 9 km resolution domains. For the rest of the finer resolution domains no parameterizations are used. Grell's cumulus parameterization is based on the rate of destabilization, essentially a simple single cloud scheme with updraft and downdraft fluxes and compensating motion determining the heating/moistening profile. It is an appropriate parameterization scheme for this level of resolution.

3. The MRF Planetary Boundary Layer (PBL) scheme for all domains. It is an efficient scheme suitable for high resolution in PBL.
4. The Reisner (mixed-phase) explicit moisture scheme [6], in which cloud and rainwater fields and ice processes are predicted explicitly. Furthermore, it includes a more sophisticated phase change approach that allows for the existence of super cooled water and ice at temperatures slightly above freezing. The scheme has no Graupel or riming processes.
5. A cloud radiation scheme that is sophisticated enough to account for long wave and short wave interactions with explicit cloud and clear air.
6. A 5-layer soil ground temperature scheme [7]. Temperature is predicted in 1, 2, 4, 8, and 16 cm layers with fixed substrate below using the vertical diffusion equation.

Since this prediction is intended for the operators of the telescopes on Haleakala, the area of interest is the Hawaiian Islands; specifically concentrating on island of Maui. Since storm systems miles away can affect the Hawaiian Islands, the prediction must include a long range forecast. The Hawaiian Islands contain a variety of microclimates in a very small area. Some islands have rainforests within a few miles of deserts; some have 10,000+ summits only a few miles away from the coastline. Because of this, the model must be run at a much finer resolution to accurately predict these areas. To satisfy both requirements, a nested grid approach must be used. The MM5 model uses a conventional 3:1 nesting scheme for two-way interactive domains. This allows the finer resolution domains to feed data back to the coarser domains. The largest domain covers an area of approximately 5000 km by 5000 km at a 27 km grid resolution. It is then nested down to 9 km around the Hawaiian Islands and then down to 3 km for each of the 4 counties. Over Maui, the grid is nested down to 1 km over the summit of Haleakala.

4. DAILY OPERATIONS

Every Night at Midnight Hawaiian Standard Time (HST), a PERL script is run handle the entire operation necessary to produce a forecast and post it to the MHPCC web page (http://weather.mhpcc.edu/mm5). The procedures the script executes are:

1. Determine and download the latest global analysis files from NCEP for a 48-hour simulation,
2. Begin processing by sending these files through MM5's REGRID program,
3. Take the output data files from REGRID and input into INTERPF,
4. Prepare the MM5 model for the current simulation,
5. Submit the MM5 run to MHPCC's IBM SP3 (*Tempest*) for execution (daily reservation starting at 1 A.M.),
6. Average daily run requires between 7 – 7.5 hours for completion on 32 processors (2 nodes),
7. Data is output in 1-hour increments,
8. Data is processed in parallel to create useful images for meteorological examination,
9. Convert images to a web viewable format,
10. Create the web pages these images will be posted on,
11. Post web pages and images to MHPCC's web site.

Most of these stages are self explanatory, but some require additional information. Step 1, can require some time as the script is downloading 8 distinct 24 MB global analysis files from NCEP. This can affect the time it takes for the entire process to complete as the download time can vary based on the NCEP ftp site, web congestion, and MHPCC's connectivity. In addition, the data is posted to the NCEP ftp site starting at 11 P.M. (HST) and complete any time from 11:45 P.M. to 12:00 P.M. (HST). Step 5, job submission, is handled through a standing reservation for 2 nodes (32 processors) starting at 1 A.M. (HST). This ensures that the model will be run and completed at a reasonable time in the morning; unfortunately this can be overridden, however this is a rare occurrence. Step 8, data processing, includes the choices of fields to be output to the web. Current choices are: temperature, wind speed & direction, clouds (see discussion in following paragraphs), relative humidity, and rainfall. There is an additional field produced for the 1 km summit grid, a sounding (includes temperature, dewpoint temperature, and wind direction at vertical levels above the summit). A more detailed description is given below:

1. Temperature (in degrees Fahrenheit): This field provides the temperature at the lowest sigma level (.995). Sigma of .995 conforms to an Elevation of 36 meters (118 ft) above sea level and 3076 meters (10089 ft) at the Haleakala summit.

2. Wind (meters per second): The wind speed in knots is approximately twice the value (1.94 knots per m/s and 2.237 mph per m/s). Sigma of .995 conforms to an Elevation of 36 meters (118 ft) above sea level and 3076 meters (10089 ft) at the Haleakala summit.
3. Clouds (low, middle, and high levels): Clouds are plotted on sigma levels, which are terrain following coordinates as shown in Fig. 2. For this reason the low, middle, and high layers are dependent on the elevation of the surface. The range of sigma for low clouds is from 0.995 to 0.870. This corresponds to an elevation of 36 meters (118 ft) to 1093 meters (3585 ft) at sea level and an elevation of 10089 meters (89 ft) to 3912 meters (12,832 ft) at the Haleakala summit. The range of sigma for medium clouds is from 0.870 to 0.425. This corresponds to an elevation of 1093 meters (3600 ft) to 5480 meters (17,975 ft) at sea level and an elevation of 3912 meters (12,830 ft) to 7750 meters (25,430 ft) at the Haleakala summit. The range of sigma for high clouds is from 0.425 to 0.025. This corresponds to an elevation of 5480 meters (17,975 ft) to 13,681 meters (44,875 ft) at sea level and an elevation of 7750 meters (25,430 ft) to 14,019 meters (45,984 ft) at the Haleakala summit. The clouds are defined by the cloud water-mixing ratio. For the low cloud field, clouds (gray shading) are plotted at mixing ratios greater than 0.15 g/kg. For the middle and high clouds the gray shading is for mixing ratios greater than 0.125 g/kg.
4. Relative Humidity (% with respect to water): This field provides the relative humidity at the lowest sigma level (.995). Sigma of .995 conforms to an Elevation of 36 meters (118 ft) above sea level and 3076 meters (10089 ft) at the Haleakala summit.
5. Rainfall: This plot has the accumulated model rainfall over the past hour before the specified time. The rainfall is in mm (1 in = 25.4 mm).
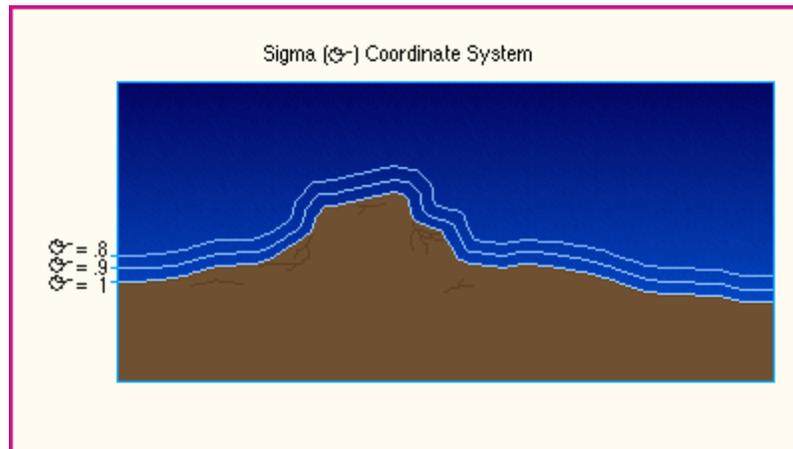

Fig. 2 Sigma terrain conforming

Additional capabilities have been added to the process of obtaining these forecasts [8]. They include:

1. Highly reliable (fault tolerant) scripts that allow for quick changes
2. Script will retrieve the most recent pre-processing data (global analysis, observational data, etc)
3. Parallel image and data post-processing for web posting

The fault tolerant script ensures that the operation will adjust and continue even in the face of an error or will report that there is a process ending error. The script has been written to be smart enough to retrieve the latest pre-processing data if it is not already present on the system; this ensures that the simulation will have the most recent data and/or avoid downloading data that is already present. Parallel image and data processing (through the use of child processes on a 16 way SMP) has been show to achieve a speedup close to the number of images/fields being produced for each domain. For example, we achieve a speedup of approximately 6 when processing 7 images (image production takes about 15 minutes in parallel relative to 90 minutes sequentially for a 48 hour simulation). This type of parallelism allows the capability of plotting more fields without significantly increasing the total image processing time.

5. WEB OUTPUT

Now that the above processes have created images, they must be made available for the telescope operators. This is accomplished by posting to the MHPCC web page; specifically http://weather.mhpcc.edu/mm5. This title page (Fig. 2) gives the user the option of what area and resolution they would like to examine.
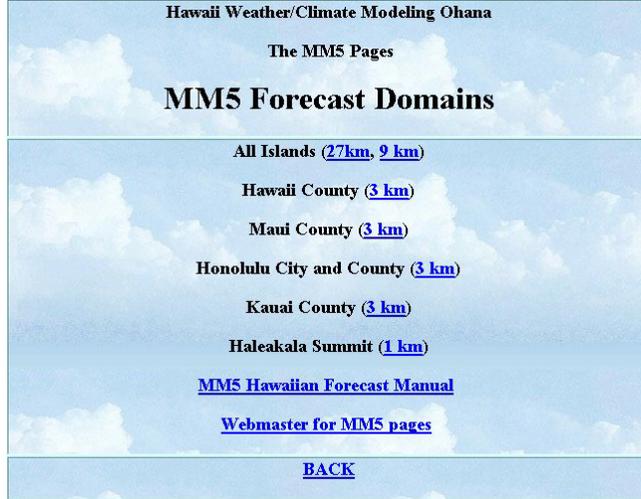

Fig. 3 Top Level of MM5 Forecast Web Pages

From the title page, the user can select the all island area at a 27 or 9 km resolution, one of the 4 counties (Hawaii, Maui, Oahu, and Kauai) at a 3 km resolution, or the summit of Haleakala at a 1 km resolution. Once one of the above has been selected, the user is transported to a web page that initially includes an image of the temperature in the selected area (Fig. 3).
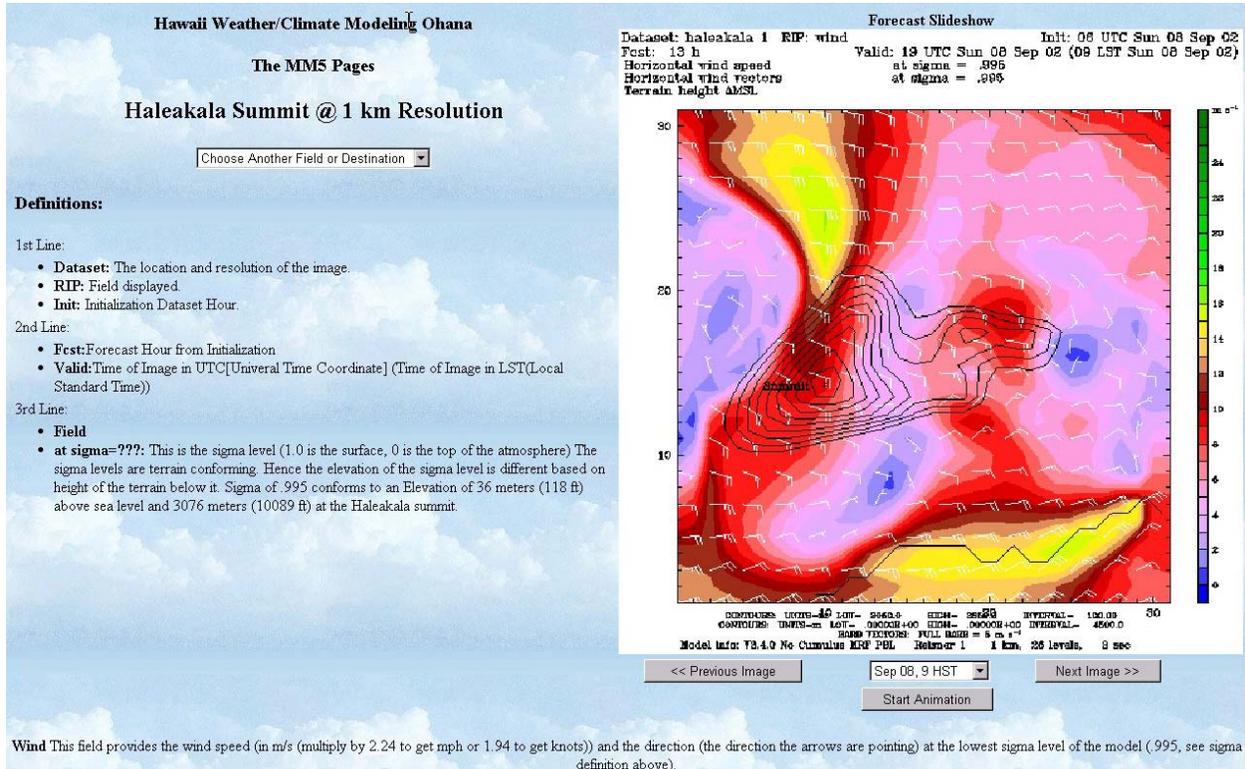

Fig. 4 Regional MM5 Forecast Web Page

On the regional web page, the viewer can select to see the previous or next image through the use of small JavaScript. If the viewer prefers, an animation of the images (in 1 hour increments) can be started and stopped. Finally, the user can select any of the images from a pull down menu. If the viewer would like to change the field being examined, a pull down menu on the left side of the page will transport the user back to the main menu or allow them to choose a different field. Lastly, if the JavaScript becomes a problem for the viewer's browser, they have the option of being switched to a non-JavaScript equivalent version web page.

6. VALIDATION

In order to ensure accuracy, some form of validation needed to be done. The most obvious choice was to use the CAPS weather sensor output from Haleakala (http://banana.ifa.hawaii.edu/caps/CAPSdata.html). These 10 sensors allow the comparison of the results of MM5 predictions against an actual measurement. Another set of sensors used for validation are those owned by the Hawaiian Commercial & Sugar (HC&S) company. They have an extensive network of sensors; unfortunately, the use of this data is limited. One can simply do a simple comparison of the sensor data to the model output at the corresponding time period. However, there are some difficulties that make validation of the model's predictions difficult. First, the model outputs on the hour (although this is configurable down to the minute but becomes impractical to do this for a 48-hour simulation) so that would need to be matched up to the sensor output only for the top of the hour. This is not a major obstacle, but it would be better to compare the output on a smaller time scale that matched the output of the sensors. Secondly, the sensor are usually less than 20 feet above the ground, the model's predictions are anywhere from 90-120 feet above the ground. This altitude difference can adversely affect the accuracy of any comparison. Lastly and most importantly, a storm appears earlier or later than the time frame the model predicted. Hence the model has predicted an event (storm, cloud cover, rain, high winds, etc.) sooner (or later) than it actually occurs. The model has still predicted the event, just not at the exact moment that it occurred. This makes validation very difficult.

To make the best attempt at validation, one needs to look at trends in the actual weather and how well the model predicts them. If the model predicts high winds from 1-3 P.M. and the high wind actually occur from 3-5 P.M. then model still has predicted the storm within a 2-hour margin of error. In addition, another method for validation can be to match up the sensor output for these events and compare to the model's predictions. If the model predicted 40 M.P.H. winds (average speed) during the event (from 1-3 P.M.) and the actual high wind were in the neighborhood of 40 M.P.H. (average speed) during the actual event then this is an acceptable validation.

The model has captured all major storm systems that have entered the state of Hawaii. Smaller, more localized events are usually captured, however the model may predict them to be slightly less/more powerful than in reality. Commonly known events such as the trade wind inversion, diurnal weather patterns, orographic rainfall, and Kona (leeward) storms are also well predicted.

7. SCHEDULING & BENCHMARK

In order to produce daily operational forecasts, a schedule must be maintained. In addition, a choice must be made on how many processors will be utilized for the model's execution. The procedure the PERL script runs through daily has been described in detail in Section 4. In addition to this, a choice was made for the number of processors to be used by the model. In order to determine this, a benchmark was done to determine the total processing time and the parallel efficiency. Processing time was examined so that we may maintain the schedule describe in Section 4. The goal of having prediction ready before 10 AM would be helpful to operators to determine a schedule for the current and following evening. The parallel efficiency was examined to determine the most cost effective choice; even though the efficiency will go down as more processors are used, the efficiency will be compared to the benefit of the total processing time for the model's completion. The benchmark for the daily Hawaiian Islands run can be seen in Table 1.

A relative speedup is used because the job had a high memory requirement that forced it to be run on no less than 16 processors. Even though this benchmark may not seem so impressive at first, this is a very complex set of nested domains. The choice of 16 processors can be dismissed, because it prevents the run from being completed by our 10 AM (HST) goal. The 32-processor relative efficiency of 84% is acceptable because it allows us to complete the model run before the 10 AM (HST) goal. If more processors are used, the model run is completed faster, but at a higher price. Since the goal is met with 32 processors, there is not need to use more processors. In the future, there may be a need to extend the forecast to more than a 48-hour simulation. At this time there may be a need to start the run earlier or to utilize more processors to complete the model run in the specified time goals.

	Time (Seconds)	Time (H:M)	Relative Speedup	Parallel Efficiency (Relative to 16 processors)
16	42790	11:53	1.00	1.00
32	25380	7:03	1.69	0.84
64	17078	4:45	2.51	0.63
128	12663	3:31	3.38	0.42

Table 1: MM5 Benchmark for the Hawaiian Islands

8. FUTURE WORK

There is plenty of additional work that can be done to improve the weather predictive capabilities. Some will help the reliability of the model in producing a forecast in the required time frame; others will help the accuracy of the model. A list of potential future work includes:

1. Porting the code to another machine. The model currently runs on MHPCC's IBM SP3; this system is heavily loaded. Currently, nightly reservations prevent congestion on the machine from interfering in the daily predictions from executing. However, there are exceptions that can prevent a run from occurring. A possible way to prevent this from happening is to have a backup machine to run the MM5 model with in case an exception occurs. MHPCC's Linux cluster is a potential candidate for this application. Once a code is ported there then a script can be setup to make sure that if a run can not occur on the IBM SP3 in the required time frame then MM5 should be run on the Linux cluster.

2. Observation data inclusion. Currently no observational data is included in the daily predictions, it is only used now for validation. This was determined to be difficult to implement, as only 1 to 2 hours of observation data may be available in the required time frame needed. If a better system can be developed to obtain the data and excluded bogus data through an automated script (i.e. the observation data will need to be available for download), then observation data can be included.

3. Better terrain and vegetation data. Currently the model uses 30-second (~0.9 km) terrain elevation and vegetation data. This appears acceptable as the lowest resolution domain is 1 km. However, the data points in the data set do not mesh completely with the domain grid points (interpolated). If a finer data set is used then the interpolation would most likely be more accurate. In addition, the vegetation data sets for the Hawaiian Islands are not completely accurate. Urban points are specified for the Honolulu area but the rest of the Islands are specified as Savanna. Inclusion of a better vegetation data set may lead to more accurate weather predictions.

4. Increased vertical resolution. Currently 26 sigma levels are used, with a clustering near the surface. Studies are being conducted to determine if additional levels improve the model predictive capabilities. If they prove to be worthwhile, the increase in vertical levels will correspond to a linear increase in execution time. Hence, if the vertical levels are doubled the execution time will approximately double. This may force the prediction to be done on more processors in order to meet scheduling requirements.

5. Increase horizontal resolution from 1 kilometer to sub-kilometer. This is dependent on the previous future work of better terrain and vegetation data. Since 30-second data is roughly equivalent to 0.9 km, a finer data set is required to even attempt sub-kilometer MM5 runs. In addition, more research must be done to investigate the accuracy of model at this resolution. It is not entirely clear how the model will behave at a finer resolution and hence the physics packages used by the model may need to be improved and/or modified.

6. Attempt the same forecasts with the next step in weather modeling codes. The Weather Research & Forecasting (WRF) model will be the next generation of numerical weather modeling that will eventually replace MM5. The WRF model is still under development and will not be ready for another 2+ years; However, early beta releases are being tested and it would be useful to determine how effective the model will be for the high resolution predictive qualities the Hawaiian Island forecasts require.

9. SUMMARY

A methodology has been created that will produce fine resolution weather forecasts for the state of Hawaii. This methodology is focused on providing the required forecasts in a minimal time as to still be useful to everyone from the general public to scientist at Haleakala using it to determine if the weather predictions are within their operational limits. The web output has been chosen to given telescope operators the necessary fields needed to make decisions on if weather conditions will allow the utilization of the telescopes on Haleakala. This will allow better scheduling and improve the potential efficiency of telescope operations.

10. REFERENCES

1. Michalakes, J. "The Same-Source Parallel MM5. Journal of Scientific Programming." **8**, 5-12. 2000.
2. Michalakes, J., Chen, S., Dudhia, J., L. Hart, J. Klemp, J. Middlecoff, and W. Skamarock: "Development of a Next Generation Regional Weather Research and Forecast Model." *Developments in Teracomputing*: Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology. Eds. Walter Zwieflhofer and Norbert Kreitz. World Scientific, Singapore. pp. 269-276. 2000.
3. Dudhia, J. "A nonhydrostatic version of the Penn Stat/NCAR mesoscale model: Validation test and simulation of an Atlantic cyclone and cold front." *Mon. Wea. Rev.*, **121**, 1493-1513. 1993.
4. Chen, Yi-Leng and Feng, J., "Numerical Simulation of Airflow and Cloud Distributions over the Windward Side of the Island of Hawaii. Part I: The Effects of Trade Wind Inversion." *American Meteorological Society*, 1117-1134. May 2001.
5. Chen, Yi-Leng and Feng, J., "Numerical Simulation of Airflow and Cloud Distributions over the Windward Side of the Island of Hawaii. Part II: Nocturnal Flow Regime." *American Meteorological Society*, 1135-1147. May 2001.
6. Reisner, J., Rasmussen, R.J., Bruintjes, R.T., "Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model." *Quart. J. Roy. Meteor. Soc.*, **124B**, 1071-1107.
7. Dudhia, J. "A multi-layer soil temperature model for MM5." *The 6th PSU/NCAR Mesoscale Model Users' Workshop*, 1996.
8. Roe, K.P., Stevens, D., "High Resolution Weather Modeling in the State of Hawaii," *The Eleventh PSU/NCAR Mesoscale Model User' Workshop*, Boulder, Colorado, 2001.