

in the model map transformation and boundary condition codes, inadequacies in the initialization of moisture over and near the Gulf of Mexico, and the need to include a convective parameterization scheme in the model to avoid systematic thickness errors. An interesting use of the temporal behavior in model forecast error statistics was shown to be uniquely capable of revealing certain systematic errors.

Andrew Staniforth described an international community effort in the verification of mesoscale models--the COMPARE (Comparison of Model Prediction and Research Experiments) Project under the auspices of the CAS/JSC Working Group on Numerical Experimentation. Recherche en Prevision Numerique of Environment Canada has taken the lead in this effort. The idea is to compare in a controlled way the results of many mesoscale models (and eventually mesoscale data assimilation systems) from both research and operational communities, on a representative testbed of cases selected primarily from Intensive Observation Periods of well-instrumented observational campaigns. Through model and data assimilation intercomparisons, they hope to improve communication among modelers, increase understanding of mesoscale phenomena and predictability, and improve the performance of various components of mesoscale models (particularly parameterizations) and mesoscale data assimilation systems, leading ultimately to improved models for both operational forecasting and research applications. The first case chosen for this project is the March 6-8, 1986 IOP of the Canadian/US CASP/GALE (Canadian Atlantic Storms Project/Genesis of Atlantic Lows Experiment) field study. The common data set (stemming from a regional reanalysis of the observed data, and using a "standardized" data-distribution format to facilitate the distribution of future cases) has recently been distributed to participants to initialize their models and make forecasts. A workshop is planned for the Spring of 1994 to discuss the ensuing results. The preparation of a second case, an IOP drawn from the Franco-Spanish PYREX field study, is underway, and the selection of further cases is under discussion. *The chairman of COMPARE ended his presentation by inviting the organizers of the CME to consider collaborating with the COMPARE project by providing a set of scientific hypotheses to be verified by coordinated numerical experimentation using data and analyses drawn from a CME IOP.*

N 9 4 - 2 4 3 9 2

3.11 Techniques and resources for storm-scale numerical weather prediction

Kelvin Droegemeier: Non-hydrostatic modeling and prediction in the mid-1990's and beyond: strategies for dealing with data, parallel processors, and communication

Georg Grell: Multi-scale application of the 5th-generation PSU/NCAR mesoscale model

- James Doyle: The coupling of non-hydrostatic atmospheric and hydrostatic ocean models for air-sea interaction studies*
- Su-Tzai Soong: A numerical simulation of cloud formation over complex topography*
- William Skamarock: Adaptive grid simulations of convection*
- David Bacon: An unstructured grid, non-hydrostatic meso/cloud scale model*
- Andrew Staniforth: Efficient mesoscale modeling for multiple scales using variable resolution*
- Andrew Crook: Initialization of cloud-scale models with Doppler radar data*
- Robert Wilhelmson: Making effective use of future computing architectures, networks, and visualization software*

The keynote presentation by Kelvin Droegemeier outlined the principal challenges associated with operational storm-scale prediction. With regard to data access, the realtime collection and processing of WSR-88D data will require high-bandwidth networks and relatively sophisticated database management systems; little work has been performed to date in either of these areas. Numerical models continue to evolve and mature, and adaptive grid refinement appears to show great promise for providing selectively enhanced spatial resolution in critical regions of the flow. *The major areas of difficulty in modeling center around accurate characterization of surface features and processes and obtaining quantitative information on water substance fields.* Data assimilation methods are receiving considerable attention in the scientific community, and the adjoint technique, though still relatively expensive, appears suitable for application to moist processes (e.g., the tangent linear approach), though application to storm-scale prediction has not yet been attempted. Finally, although massively parallel computers continue to be made available, the realized performance for most codes is far below machine peak due to the absence of suitable translation tools. The advent of Fortran-P and High Performance Fortran should help alleviate this problem, though it will likely be some time before production scientific experiments are run on MPP's on a regular basis.

Georg Grell introduced the fifth generation of the PSU/NCAR mesoscale model. The MM5 model allows for the integration of the hydrostatic as well as the non-hydrostatic equations of motion. The physics routines include many different choices for implicit treatment of convection, a mixed phase explicit treatment of cloudwater, rainwater, snow, and ice, and a choice of boundary layer parameterizations including the Blackadar PBL scheme as well as a second order Mellor-Yamada closure scheme. The mesh refinement scheme allows for an unlimited number of nests and nest-levels, which can translate and overlap. An example of a multi-scale nonhydrostatic application was given for an explosively deepening ocean storm case from ERICA. For this case, the resolution of the two-way interactive grids covered four levels (75, 25, 8.3, and 2.8 km). The successful simulation used two stationary domains (at 75 and 25 km resolution) and three translating

domains, which were moving with either a cold front (8.3 km resolution domain) or the center of the surface storm (8.3 and 2.8 km resolution domain). The model was able to capture the spin-up and overall structure of this event, as compared to satellite observations. The MM5 code is now being moved to a massively parallel system (the Intel) at Argonne, as well as to a cluster of IBM RS6000 workstations.

James Doyle showed results from a three-dimensional coupled ocean/atmosphere mesoscale prediction system (COAMPS). This model consists of nonhydrostatic atmospheric and hydrostatic ocean models, which can be integrated simultaneously. Idealized numerical simulations of Hurricane Gilbert demonstrate the capability of COAMPS to realistically simulate the mesoscale air-sea interaction processes. Mesoscale atmospheric structures such as spiral bands of convection and concentric eye walls developed in the simulations. In the coupled simulations, the intensity of the hurricane was hindered by strong cooling of the sea-surface temperatures forced by the the hurricane circulation. Other numerical simulations that were presented include a convective event and a case of marine atmospheric boundary boundary layer frontogenesis.

Su-Tzai Soong reported on a non-hydrostatic, fully compressible mesoscale atmospheric convection model which was developed to study heavy precipitation induced by the complex topography over southwestern Taiwan. The unique feature of this model is the adaptation of a third order advective scheme which preserves the peak values well and produces no phase error and little computational oscillations. The steep and complex topography over Taiwan can generate such oscillations, which if not dealt with properly, may either generate false clouds or cause computational instability. The scheme demonstrated no undesirable oscillation in a simulation of a thermal using a two-dimensional dry version of the model without applying any smoothing. The same model was used to simulate many intricate features of mountain waves without producing any ill effect. The complete two-dimensional model, including microphysical processes and radiation, was used to study the effect of complex topography in producing heavy precipitation in southwestern Taiwan produced by a rainband of the decaying typhoon Agnes, which passed the east coast of Taiwan three days before. The simulation clearly showed the formation of a storm over the mountain area and subsequent propagation of the storm to the southwestern plain caused by the continuous generation of new cloud cells at the western edge of the storm. The westward propagation of the storm combined with the eastward propagation of cells caused heavy precipitation and flood over the southwestern plain.

William Skamarock illustrated the use of adaptive grid refinement for a three-dimensional simulation of an intense squall line/MCS using the COMMAS model. A video of the simulation showed how the various sub-meshes were arranged automatically through an algorithm based on the location of the updraft maxima in the domain. This type of gridding strategy allows one to capture the details of convective evolution while also representing the environment in a credible way.

David Bacon introduced a novel approach for cloud/mesoscale modelling. The Operational Mesoscale Environment model with Grid Adaptivity (OMEGA) is a model built on a triangular prism mesh which is structured vertically, but is horizontally unstructured. The mesh is capable of continuous refinement so that OMEGA has variable horizontal grid resolution ranging from 100 km down to 1 km. A key feature of OMEGA is its inclusion of a wide variety of terrain, land/water, land use, and vegetation data. The other major advance of OMEGA, slated for the 1993-1994, will be the dynamic adaptation of the grid. By letting the grid resolve the regions of severe weather, OMEGA will achieve previously unachievable resolution in a mesoscale model.

Andrew Staniforth presented the principal attributes of a highly-efficient, variable-resolution, nonhydrostatic global model currently under development in Canada. This model is designed to be easily reconfigurable for multiscale applications ranging from the global all the way down to the meso-gamma scale. The numerical techniques include a finite-element spatial discretization (for variable resolution) on a rotated latitude/longitude mesh (to focus resolution anywhere over the globe), and a semi-implicit semi-Lagrangian time integration scheme (to avoid an overly-restrictive timestep limitation). Because the model is global, problems due to the imposition of lateral boundary conditions and their specification are avoided. Results from a shallow-water prototype were given to illustrate the potential flexibility of this approach to handle multiscale problems. A particularly interesting and surprising property of the horizontal mesh used is that fully 50% of the total number of meshpoints are contained within a 1/4 km uniform-resolution square sub-domain of length 100 km, yet the mesh is global in extent and each successive meshlength is only 10% larger than its preceding neighbor as one moves away from the high-resolution window. It was shown that a mesogamma-scale feature embedded in a synoptic flow can be successfully retained and forecast without deteriorating the steering synoptic and global-scale flow. If this behavior holds up for the three-dimensional model under development, then it was argued that it could prove to be a useful tool for multiscale research in the context of the CME.

Andrew Crook spoke on the problem of initializing cloud-scale models with single Doppler radar data. Two approaches to the problem were presented. In the first, which is mainly applicable to boundary layer flow, both components of the horizontal wind are determined by a tracking technique called TREC (Tracking Reflectivity Echoes by Correlation). The pressure and buoyancy can then be retrieved from a time history of the horizontal wind and application of the continuity equation to obtain the vertical wind component. Tests with this method on a number of gust front cases have shown that the motion of the front can be predicted with reasonable accuracy. The second approach to radar data initialization is the adjoint technique. The method attempts to find the initial conditions that produce a simulation which fits the available data in a least squares sense. The technique has been applied to single Doppler data of a gust front case from the Phoenix-II experiment. The horizontal velocity field has been retrieved to within 1.5 m s^{-1} rms error when compared with dual Doppler data.

Robert Wilhelmson spoke on tools currently available, as well as those planned for the next several years, for application to large-scale computational problems. Emphasis was placed on homogeneous computing environments across the spectrum of devices (mass storage, compute engines, visualization environments), and on increased resolution for visualization devices. No longer are supercomputers found only in government sponsored facilities such as NCAR, GFDL, and NWS. Today, there are many atmospheric scientists who use these new supercomputing facilities that are available to remote users over the Internet, a national network interconnecting hundreds of local and wide-area networks. This network has a cross-country backbone with network speeds between nodes at or above 1.5 megabits per second. This is more than 1,000 times greater than the 1,200 bits per second available during the early 80's. The next decade will be filled with new capabilities for researchers to carry out their investigations. For example, in the latter half of the decade a teraflop computer should be available. Computers that deliver this speed will most likely be highly parallel, with over 100 processors. Some of today's massively parallel machines already have over 1,000 processors. *The key to obtaining performance on these machines is the efficient use of these processors.* This has led to a growing interest in algorithms that can be used to achieve this high efficiency. Today's conventional gigaflop supercomputing will be handled in the near future by computers that can be purchased by a research group and placed in an office environment. If several of these computers are available, they can be interconnected through high speed links and simultaneously used to do a large calculation. Further, these computers can reside on high speed networks interconnecting supercomputers and personal computers. Larger memories will be coupled to increased computer speed, with maximum configurations well beyond the 32 billion byte memories

available today. In addition, the handling and storage of data will be expedited by larger disk farms, the use of new storage media, and higher speed access. DD-2 tapes, widely used in the professional video marketplace, can hold over 150 gigabytes of data (per tape), almost 1,000 times more than the archive tape in current use at many supercomputer sites.

Computers, their interconnections, and data storage systems are being referred to as a 'metacomputer.' The metacomputer is a natural extension of parallel computers in which all processors are alike. Interconnections between different processors can reach over one gigabyte today and these speeds should increase significantly through the use of parallel channels over the decade. Within the metacomputer there will be facilities for making video tapes. High definition technologies will enable creation of these videos with substantially more resolution, at least as high as today's typical workstation monitors (approximately 1000 x 1000 pixel resolution). Sound will be integrated with visual images to provide another information dimension. Virtual reality environments will be used to transport the researcher into their computational space, experiencing three dimensionality and integrating their data based on this three dimensional stereo view. For example, it will be possible to walk under a growing storm and release air tracers or to fly through a storm checking for regions of high turbulence. The metacomputer is also an integrating concept for access to a wide variety of data needed to solve problems, often interdisciplinary in character, where small and large specialized data sets are stored at many sites throughout the country. Data management systems are needed by the researcher to check on what data is available and then to obtain portions of it over the national network. Further, the metacomputer will be used by groups or individuals around the country carrying out collaborative research. With new technological advances it will be possible for researchers at remote sites to collaborate as if they were in the same room, seeing the same computer screen, talking and pointing to what they see, using the computer as a blackboard, developing code together, and jointly controlling and monitoring simulations and data analysis. This will be important for tackling grand challenge problems such as the study of global change and the improvement in mesoscale weather forecasting.

Advances in computer technology and its use is being fostered through The Federal High Performance Computing and Communications (HPCC) Program. The strategic priorities of this program are to further U.S. leadership in HPCC, to increase industrial competitiveness through the use of HPCC technologies, and to accelerate the widespread application of these technologies for economic, national security, education, and global environmental purposes. The atmospheric science community is benefiting significantly from participating in these efforts, with new computational environments being explored.

Wilhelmson presented an example of a computational environment of the 90's called PATHFINDER (Probing ATmospHeric Flows in an INteractive and Distributed EnviRonment), a project at the National Center for Supercomputing Applications and NASA Goddard Space Flight Center. The mission of the PATHFINDER Project is to create a flexible, modular, and distributed environment for data handling, model simulations, data analysis, and presentation to be used in studying atmospheric and fluid flows, and which can be tailored for specific scientific research and weather forecasting needs.

4. Workshop proceedings and recommendations

Following the colloquium, the participants met in workshop breakout sessions to use the previous discussions as a base for discussions. The workshop session leaders were assigned five tasks by the meeting organizer to be completed before the end of the meeting:

- Assess the advantages and limitations of current modeling approaches being used in the topic area covered by that particular session
- Define the critical unknowns in the topic area
- Make specific suggestions for future modeling approaches, with emphasis on those activities which should be accomplished prior to the 1995 CME
- Determine the implications of these developments and approaches as they impact future observing system strategies
- Determine the observational requirements for coupled multiscale model validation

4.1 Joint sessions on initialization and data assimilation

Tom Warner and Fred Carr

a. Advantages and limitations of current approaches

Intermittent data assimilation

Concerning the characteristics of the intermittent four-dimensional assimilation approach first, the workshop made the following conclusions:

- The main *advantage* of this approach is that it is simple and uses primarily existing software (i.e., the model and objective analysis programs).

However, there are several *disadvantages* to this approach:

- There can be considerable model "shock" resulting from the inaccurate dynamic balance that results from the intermittent objective analyses.
- Associated with the dynamic imbalance is the lack of accurate mesoscale vertical motions resulting from each objective analysis; thus, a significant fraction of each forecast interval may be required in order for reasonable vertical motions to be generated internally by the model. As the assimilation window (model restart interval) is reduced, the model has less time to recover from the initialization shock and to spin up mesoscale structures such as vertical motion fields.
- The objective analysis of often primarily synoptic-scale data can remove mesoscale structures generated during the model forecast through local forcing, nonlinear interactions, etc.
- It is difficult to utilize "nonstandard" data that cannot be analyzed as point values of the predicted variables.
- The model initial state is not "optimal" in any sense.

Continuous data assimilation

Concerning the popular continuous data assimilation ("nudging") approach, the workshop concluded that the *advantages* of this approach are:

- The software is relatively simple to implement.
- It is economical to use relative to many of the next-generation variational procedures.

The *disadvantages* of this approach are somewhat similar in nature to those of the intermittent approach, namely:

- Model shock may result when the relaxation terms are shut off in data voids (in space and time).
- It is difficult to utilize "nonstandard" data such as rain rates and satellite cloud images.
- The model initial state is not "optimal" in any sense.
- Although the computational cost is much less than that of the next-generation variational procedures, continuous assimilation can become computationally costly if many data are available.

Physical initialization

The *primary advantage* of the physical initialization technique is that:

- The data utilized generally have mesoscale resolution (e.g., latent-heating rates based on radar-reflectivity), and thus their use can reduce, or at least not inhibit, model spinup time.

The *disadvantages* of this approach are:

- It is often difficult to quantitatively interpret the data used in this approach because they are based on indirect measurement systems (e.g., radars and satellite sensors).

- This approach is often model specific because it is sometimes closely tied to physical-process parameterizations.

Variational approaches

The *advantages* of the variational approach are:

- Observations of both "traditional" (u,v,T) and "non traditional" (e.g. - radiances, reflectivity) variables can be incorporated into the assimilation.
- The cost function can include observations, background fields (usually prior forecasts), as well as additional dynamical and/or smoothness constraints to accomplish the desired goals of the analysis.
- The technique is adaptable for any type of model and scale; if the model equations/physics are appropriate for a particular phenomena and the data are "sufficient," then an optimal solution can be found.

The *disadvantages and unresolved issues* with respect to variational techniques also represent the critical areas for research:

- *Discontinuous processes*: The treatment of switch on/off physics such as cumulus parameterization causes difficulty in the minimization; this is an active area of research and some promising solutions have already been identified.
- *Computational expense*: Ten or more iterations of forecast and adjoint models over the assimilation period are usually required. The use of preconditioning, lower-resolution models, simpler models/adjoints during assimilation, better minimization algorithms and faster computers may all contribute to the solution of this problem. The key operational issue is whether the extra effort and expense of this approach is worth the benefit it may provide to the subsequent forecast.
- *Model error*: The model is usually assumed perfect when in fact it has error (which may be considerable if the model physics can't simulate phenomena which are present in the data). In addition to improving the model, possible solutions include adding a model bias correction term, or making use of the Kalman filter technique, in which the model error is known at all times.
- *Error statistics*: The observational and background error covariance statistics (as well as the weights to be assigned to the various additional constraints as a function of space and time) are important to the quality of the analysis and need to be better simulated. For mesoscale and convective scale flows, the statistics are not known at all and are probably non-Gaussian. Much effort is required to successfully address these issues.

Cloud model initialization

The *advantages* of this exploratory technique include:

- Modified forward insertion and dynamic relaxation techniques have shown some success if thermodynamic retrieval methods and/or other dynamical constraints are incorporated.
- Multi-parameter radar data can be used in the initialization of hydrometeors and other cloud microphysical parameters.

The present *limitations* of this technique consist of the following problems:

- Simple insertion or nudging of wind data will not recover the correct temperature field.
- If data from only one Doppler radar is available, then single-Doppler retrieval methods need to be applied. Velocity azimuth display and related methods just provide a vertical sounding of the horizontal wind for a volume around the radar. TREC techniques apparently are restricted to clear-air boundary layer applications.

b. Critical issues

The unresolved issues and challenges in the areas of *current and next-generation data assimilation* include:

- "Optimal" nudging approaches that employ variational techniques to define relaxation coefficients need to be further developed and tested to determine their usefulness.
- The procedure needs to be further tested with mesoalpha- and mesobeta-scale models
- Specific critical issues for variational techniques were listed immediately above. The key issue can be summarized as: Even if solutions are found for most of the current problems, will the extra effort and expense of the variational methods be repaid by a significant improvement in the assimilated state and resulting forecasts?
- It is necessary to determine how relaxation coefficients can be specified in such a way that information about short-time-scale processes can be introduced without the Newtonian terms dominating the solution determined by the model dynamics.
- Concerning physical initialization, the quality and interpretation of indirectly measured precipitation data need to be improved, in conjunction with the development of techniques to retrieve cloud water, cloud ice, and other hydrometeors.
- Precipitation and hydrometeor estimates from the national network of WSR-88D radars need to be incorporated into physical initialization procedures.
- The proper vertical distribution of specified heating and moistening rates for different mesoscale phenomena needs to be determined. The role and importance of cloud microphysical processes in this problem also needs to be studied.
- Techniques need to be developed to insert information about fluxes, transports, and other process rates (e.g., evapotranspiration, rainfall rates, and TOA radiative fluxes) into analyses in a consistent fashion into coupled models.
- Another issue is how adjoint and other variational techniques can be used in coupled models such as coupled hydrological/atmospheric models.
- Little is understood about how shortcomings in model physical parameterizations and numerics affect four-dimensional data assimilation fields.

- One of the major suggestions of this workshop is the need to develop techniques to adapt present data assimilation approaches to multiply-nested or adaptive grids.

The primary problems, issues, and research ideas *concerning convective-scale model initialization* which arose from the discussions in this workshop breakout session concern data types and frequency, adjoint techniques, boundary conditions, data gaps, and initialization of microphysical processes:

- *Many questions remain on the types of data needed, the spatial density and temporal frequency required and the amount of error the techniques will tolerate.* The data frequency issue is especially critical because model time steps for cloud models are much smaller than the radar volume scan interval plus the data often have to be smoothed to reduce noise in the time tendencies; thus the assimilation window may be too large for rapidly evolving phenomena. Although retrieval techniques using wind data alone have shown some success, the presence of observed temperature data helps significantly. Knowledge of the water vapor, liquid, ice and hydrometeor fields is crucial for thunderstorm prediction. Many more real-data experiments to explore these and other issues are needed.
- Although they have shown success in the retrieval/analysis problem, *much work remains to incorporate adjoint methods into a data assimilation system for real time prediction of storm-scale flows.* All the problems listed for variational techniques above apply; the issue of discontinuous processes is even more acute here. In addition, the lack of predictability of the model for observed convective phenomena may lead to multi-minima in the cost function.
- The need for accurate boundary condition information is crucial for analysis and forecasting over small regional domains. *The treatment of boundary values in the adjoint approaches is especially important.*
- Problems exist for *specifying the initial conditions in regions where no data are available;* e.g. in the clear air environment surrounding a convective storm.
- Studies on the use of multi-parameter radar to accomplish the *initialization of microphysical processes* should be continued.

c. Observational requirements for improved data assimilation systems

Knowledge of the three-dimensional mesoscale structure of the moisture field at frequent time intervals is the most crucial advance needed. No existing observing system is capable of satisfying this requirement. The strengths of many individual systems need to be combined to overcome the current deficiencies in moisture data measurement. Important components include:

- Combined visible, infrared, and microwave channel data from satellites to infer water vapor distribution, precipitable water, cloud coverage and precipitation estimates, and cloud water, ice and hydrometeor information.
- Ground-based remote sensing techniques for moisture, especially the promising DIAL system.

- WSR-88D reflectivity data to infer surface precipitation rates as well as cloud water/ice distribution. Dual polarization and/or multi-parameter radars should be utilized to improve information on hydrometeor content and other cloud microphysical parameters.
- In situ data from conventional radiosonde measurements is very important, not only for their detailed vertical profiles of moisture but for calibration and assessment of the remote sensors.
- The CASH component of the ACARS system should be strongly encouraged; high-frequency measurements during takeoffs and landings are especially desired.
- High-resolution surface moisture measurements are important for the convective initiation problem.
- Research aircraft data will be needed in clouds to measure cloud microphysical parameters; these data are needed to initialize cloud models as well as to assess the accuracy of retrieval techniques used to deduce moisture variables from Doppler radar data.

Other important observation needs for data assimilation are:

- Co-location of frequent rawinsonde releases near remote-based sounding systems is recommended so that the relative accuracy of their velocity and temperature measurements can be assessed. Use of CLASS and mobile CLASS systems is desired. Access to the raw sounding data is requested to maximize use of the vertical detail present in the data.
- High-density raingauge networks are needed to help calibrate WSR-88D precipitation algorithms at each site. Both the gauge and radar precipitation estimates should be available to operational centers at hourly intervals.
- Soil moisture, snow cover, albedo, and other surface characteristics need to be measured to allow testing of their incorporation into data assimilation schemes. Measurements of surface fluxes are needed for verification of model estimates.
- The ACARS program should be expanded to include more aircraft and greater frequency of measurement.
- All WSR-88D radars in the CME should have recorders, not just the ones operated by NOAA.
- In order for data assimilators to make proper use of the data, the data from each observing system should be accompanied by complete documentation on the system's operating principles, measurement strategy, and accuracy, as well as on the quality control procedures used in processing the data.
- Observing system simulation experiments for future observing systems (alone and in combination with others) should be conducted to help determine what data types, accuracy, density and distribution are most important to analysis and forecast systems; this will help determine priorities for selecting which systems to design and implement in future field experiments.