BOREAL ENVIRONMENT RESEARCH 7: 193–201 ISSN 1239-6095
Helsinki 2 October 2002 © 2002

The BALTEX regional reanalysis project

Carl Fortelius[1)], Ulf Andræ[2)] and Mikael Forsblom[1)]

[1)] *Finnish Meteorological Institute, P.O. Box 503, FIN-00101 Helsinki, Finland (e-mail carl.fortelius@fmi.fi)*
[2)] *Swedish Meteorological and Hydrological Institute, SE-601 76 Norrköping, Sweden (e-mail: ulf.andrae@smhi.se)*

Fortelius, C., Andræ, U. & Forsblom, M. 2002: The BALTEX regional reanalysis project. — *Boreal Env. Res.* 7: 193–201. ISSN 1239-6095

The BALTEX regional reassimilation project uses meteorological data assimilation for quantifying the climatic energy and water cycles over the catchment basin of the Baltic Sea during the course of one annual cycle, Sep. 1999–Oct. 2000. This report presents the data assimilation system used, the available products, and a sample of preliminary results. The latter demonstrate that the system is capable of simulating the essential features of the energy and water cycles of the Baltic drainage basin. We find this encouraging, because the model has not been tuned to reproduce these cycles, but mainly to predict the atmospheric state.

Introduction

Numerical simulation of the atmosphere has a long history in weather prediction and climate research (Wiin-Nielsen 1991). Models are probably as important as measurements, when it comes to forming our present understanding of the processes which drive the general circulation of the atmosphere and oceans, and ultimately control the climate (Browning and Gurney 1999). The atmospheric models (AMs) used are based on the equations of fluid mechanics discretized in time and space (e.g. Haltiner and Williams 1980). The equations express the conservation of dry mass, energy, momentum, and water in all its phases for a parcel of air. Many processes and phenomena directly related to the energy and water cycles are not described by the conservation equations, or occur on too small scales to be resolved by the discretized equations. Undescribed processes include radiative transfer, cloud micro physics, and atmosphere-surface interactions. Convection and turbulence are examples of important unresolved processes. All these processes and phenomena are treated as source functions in the discretized equations and expressed approximately in terms of the resolved variables. The procedure is referred to as parameterization, and constitutes a major part of the effort of developing an atmospheric model. Unfortunately direct observations of the relevant processes are scarce. Therefore formulations are often judged by how well the model is able to predict observable state variables. The drawback of this method is the difficulty in relating errors in the state variables to a particular process. Sometimes errors in one process may even be masked by compensating errors in another process.

Despite their impressive veracity and accuracy, AMs are theoretical constructs, and as such they must constantly be tested against observations. Because of internal variability, long-term climate simulations can only be tested against long-term climate data. Weather forecasts, on the other hand, can be verified individually, and thus constitute a very stringent test on the model used. The possibility of direct intercomparison with available observations make simulation systems operated in forecast mode an attractive data source for climate studies.

The initial conditions for numerical weather prediction are prepared by data assimilation. This is a successive analysis of observations with the help of a dynamical model, which fills data gaps by advancing information from previous observations in time. Data assimilation is used also to create archives of the physical state history of the climate system, based on historical observations. Typically such archives include processes and phenomena related to the climatic energy and water cycles, such as precipitation, evaporation, runoff, clouds and other moisture fields, surface heat fluxes by turbulence, as well as radiation at the surface and the outer boundary of the atmosphere. In addition, the apparent net atmospheric sources and sinks of energy and water may be computed as residuals in the relevant budget equations, using the analyses to evaluate all other terms (Yanai *et al*. 1973). It must be borne in mind, however, that whichever method is applied, the products of an assimilation system are always susceptible to both systematic and random errors of the models, and should be used with care (e.g. Fortelius 1995). The analyses will to some extent be model-dependent and products should be intercompared across models whenever possible. Improvements to the resolution or formulation of models should lead to better analyses and so periodic re-analyses can be useful.

Historical data analysis employing global assimilation systems and spanning many decades has been undertaken by several research centres, such as the National Centers for Environmental Prediction (NCEP) and The National Center for Atmospheric Research, (NCAR) in the United States (Kalnay *et al*. 1996) and by the European Centre for Medium-Range Weather Forecasts (ECMWF, Gibson *et al*. 1999, Simmons and Gibson 2000).

These global data sets have too coarse spatial resolution for resolving topographic features on scales of tens of kilometers, such as many river catchments, large lakes or even the bays of the Baltic Sea. As a complement to the global systems, regional high resolution data assimilation for climate studies is therefore under way at NCEP for North America, and within the European HIRLAM-community for the BALTEX-area (Raschke *et al*. 2001). These projects use data assimilation systems based on Limited Area Models (LAMs). A LAM is an atmospheric model applied to only a part of the globe. Unlike a global model, a LAM requires externally specified boundary conditions at the lateral boundaries of its domain. These typically consist of predicted or analyzed data from a global system or from another LAM. LAMs are used extensively in short range weather prediction and for dynamical down-scaling of global climate simulations.

The present paper describes the on-going BALTEX regional reanalysis project.

Project organization and the assimilation system

The BALTEX regional reanalysis project is a joint undertaking of the Finnish (FMI) and Swedish (SMHI) national meteorological services, organized as an ECMWF special project. Its objective is to carry out high resolution data assimilation around the Baltic drainage basin (Fig. 1) over one year (Sep. 1999–Oct. 2000) during the BALTEX main experiment BRIDGE (Bengtsson 1998), and thereby to promote the use of assimilation products in regional climate system research. A specific objective is to produce gridded fields of all components needed to close the energy and water cycles, with a spatial resolution of approximately 22 km and a temporal resolution of six hours. The assessment of these products using available independent measurements constitutes a significant part of the project. Details about the project status and data archive are given on the project home page, which is linked to the BALTEX web site http://w3.gkss.de/baltex/.

The assimilation system used is a specially designed version of the HIRLAM numerical weather prediction system maintained by the

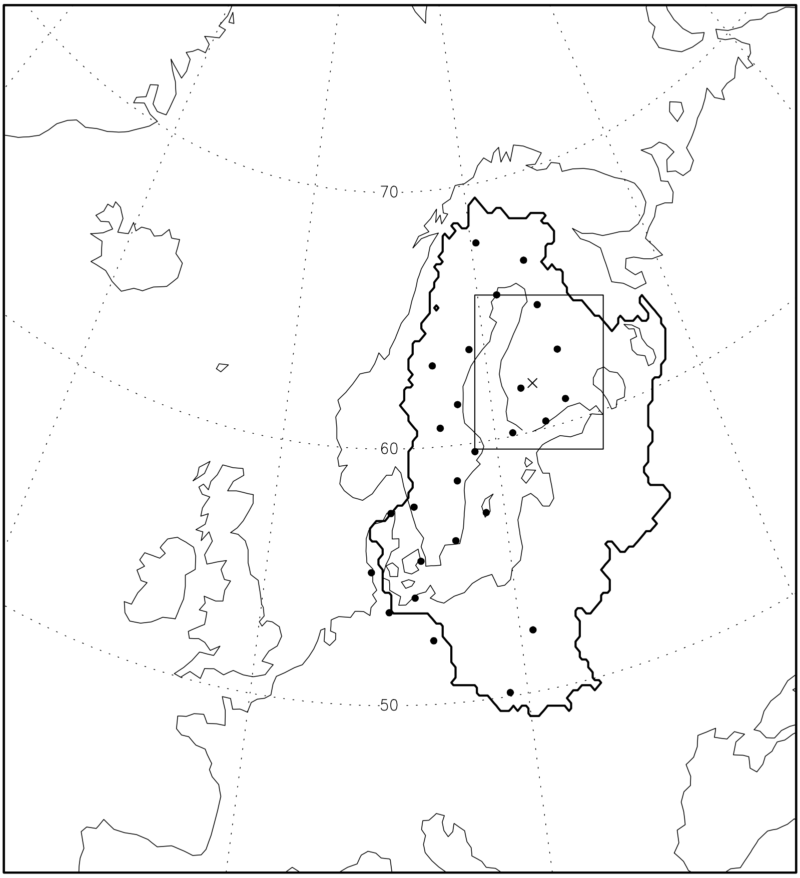

Fig. 1. The BALTEX reanalysis area. The heavy closed contour indicates the drainage basin of the Baltic Sea. Black circles indicate weather radars and the cross over southern Finland shows the location of the field station at Hyytiälä. The significance of the rectangle over Finland is explained in Preliminary results.

international HIRLAM project, consisting of the national meteorological services of Denmark, Finland, France, Iceland, Ireland, The Netherlands, Norway, Spain, and Sweden. In the BALTEX configuration, the atmospheric forecast model is a hydrostatic, two time level, semi-Lagrangian, semi-implicit limited area grid point model (Källen 1996). Prognostic variables are surface pressure, horizontal wind, temperature, specific humidity, specific cloud condensate and the kinetic energy of small scale turbulence (TKE). The model is run with a horizontal resolution of approximately 22 km, and has 31 levels in the vertical.

Parameterized processes include radiation (Savijärvi 1990), grid scale condensation and precipitation (Rasch and Kristjansson 1998), deep and shallow convection (Kain and Fritsch 1998) and vertical mixing by small scale turbulence according to Cuxart *et al*. (2000). The convection and turbulence are intimately connected through the TKE, that is used for triggering shallow convection and for defining the mixing length of the turbulent diffusion. The surface fluxes of momentum, sensible heat and latent heat are treated in the framework of a Monin-Obukhov surface layer with stability-dependent drag coefficients (Louis *et al*. 1981).

At the surface each grid box contains a certain fraction of land, sea or ice. Fluxes are computed individually for each surface type, and averaged over the grid box. As described in Bringfelt *et al*. (2001), the energy and water budgets of the soil and vegetation are treated as in the second version Rossby Centre climate model: The land parts of a grid box are divided in forest and open land. The evapotranspiration includes transpiration from the vegetation, evaporation of water intercepted on the canopy, and evaporation from bare soil. Parameterization of soil moisture and runoff is based on soil moisture variability functions traditional in hydrological models.

The analysis of atmospheric variables is based on a variational formulation (3DVAR). It minimizes a cost function measuring the distance between the model state and a background field and the model state and the observations, respectively (Gustafsson *et al*. 2001, Lindskog *et al*. 2001). The analyzed atmospheric state is filtered with respect to gravity waves using a diabatic digital filter to get a balanced initial field for the prognostic model.

The observations used are extracted from the archives of the ECMWF. They consist of surface data from reporting weather stations, ships and drifting buoys, and upper air data from radio soundings and reporting aircraft. During BRIDGE SMHI and FMI have increased the number of globally-disseminated surface observations (SYNOP-reports).

The analysis of surface variables includes assimilation of snow and sea surface temperature (SST) observations. The SST and ice evolution in the Baltic Sea is described with a coupled ice-ocean model (Gustafsson *et al*. 1998). The model SST is adjusted through a nudging process with observations from the marine service of SMHI twice a week. The numerous inland lakes in Scandinavia are described with slab and one-dimensional lake model (Ljungemyr *et al*. 1996). The lake model as well as the Baltic Sea model are forced with atmospheric data from the forecast model, and are coupled back through the updated temperature and ice fields.

The assimilation cycle is 6 h and at every cycle a 30 h forecast is run. At the lateral boundaries the atmospheric model is forced by operational analyses from the archives of ECMWF, updated every three hours.

Data archive

The BALTEX BRIDGE data assimilation archive describes the physical evolution of the climate system as analyzed and predicted by HIRLAM. Snapshots of atmospheric motion, temperature, specific humidity, specific cloud condensate, turbulent kinetic energy, diagnostic cloud cover, and surface pressure are available every six hours on the grid of the forecast model. The sea surface is described by the sea surface temperature, ice cover, and roughness, while the snow pack, soil temperature and soil water content are available over land. In addition physiographical data on orography, distribution of land and sea etc. are available, as are diagnostic variables like temperature and moisture at screen level, wind at 10 m above ground, and cloud cover.

The snapshots are augmented with the cumulative effects of parameterized physical processes. These include two-dimensional fields of radiative fluxes at the top of the atmosphere and at the surface, surface fluxes of sensible heat, latent heat, and momentum, as well as precipitation, evaporation, and local runoff. In addition, three-dimensional distributions of the flux of precipitating water, the local cumulative tendencies of temperature, humidity, and cloud condensate due to turbulence alone and to all parameterized processes together, and the temperature tendency by radiation, are stored on a sub set of the model grid, covering the catchment basin of the Baltic Sea. Other processes of interest, such as the net condensation of water vapour or the convective fluxes of sensible and latent heat can be studied by forming linear combinations of the stored fields. Like the snapshots, the cumulative fields have a temporal resolution of six hours, but a selection of two-dimensional fields is available every hour.

Preliminary results

The data assimilation for the BRIDGE year is still not completed at the time of writing, and analysis of the results is just beginning. Therefore here, we present only few preliminary results, intended to demonstrate some of the potentialities and limitations of our products.

Figure 2 shows the atmospheric water balance of the Baltic drainage basin during the autumn of 1999. Results are shown separately for the total basin and for areas of land and sea, respectively. The budgets, like all model results presented here, are based on hours six to twelve from each forecast cycle in order to avoid possible spin-up effects in the beginning of the integration.

For the land areas as well as for the basin as a whole, precipitation P exceeds evapotranspiration E all the time. The deficit is nearly compensated by convergence of the atmospheric transport, *C*, and the total amount of water remains nearly constant. Towards the end of the year *E* decreases, and *P* is nearly balanced by *C*. Over the Baltic Sea, by contrast, *P* is nearly balanced by *E*, and *C* is smaller in magnitude. *E* peaks in October, and remains high until the end of the year. The Baltic Sea remains a net source of atmospheric water until late November, but during the heavy rains in December 1999 atmospheric convergence is a significant term in the budget.

One advantage of running in data assimilation mode, relative to operating a forecast system in climate mode, without observations, is that the predicted evolution can be directly compared with the observed one. In our study we use successive analyses to estimate the true evolution. Because the model influences the analyses, agreement between the two sets does not necessarily imply that the model is correct. However, systematic differences between analyses and predictions are likely caused by errors in the latter. Comparing the predicted average rate of change of the atmospheric water content, *dQ*, to the real one, *dQ**, (as estimated from successive analyses) shows the assimilation system to be slightly out of balance ($dQ > dQ^*$). The imbalance seems to increase towards the end of the year. We will return to this point later, while discussing Fig. 4.

Figure 3 shows components of the heat balance at the surface. Results are again shown separately for the total basin and for areas of land and sea, respectively. Over land, the summer-like conditions prevailing in September are characterized by strong heating due to short wave radiation, SW, almost balanced by cooling due to long wave radiation, LW, latent heat flux *L*, and sensible heat flux *S* (in that order of importance). As the autumn progresses, SW and *L* decrease, and *S* changes sign to become a heating agent. As a result, by the end of the year winter-like conditions obtain, with LW balanced mainly by downward sensible heat flux. The heat flux into the soil, *G*, is equal to the sum SW + LW + *L* + *S*. It is relatively small, but negative nearly all the time, indicating a realistic cooling of the soil.

The sea-surface fluxes in Fig. 3 clearly show the importance of the Baltic Sea in moderating the seasonal cycle. As the autumn progresses, the release of heat stored in the sea (and lakes)

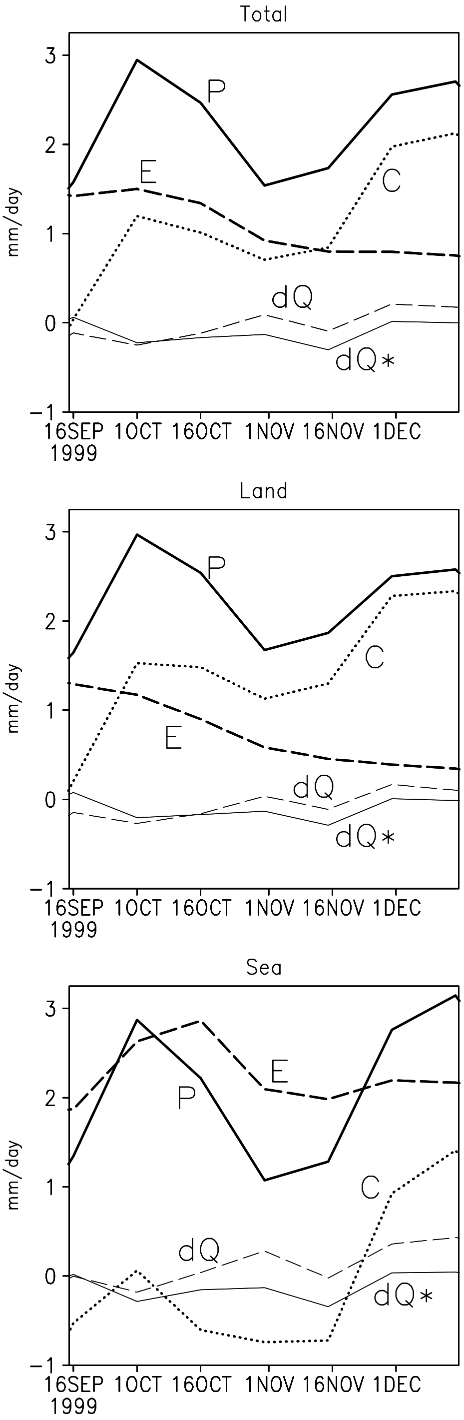

Fig. 2. Components of the mean atmospheric water balance over the catchment basin of the Baltic Sea: precipitation (*P*), evaporation (*E*), atmospheric convergence (*C*), predicted (*dQ*) and observed (*dQ**) net change in atmospheric water content. The graphs show 30-day averages in units of mm d^{-1}.

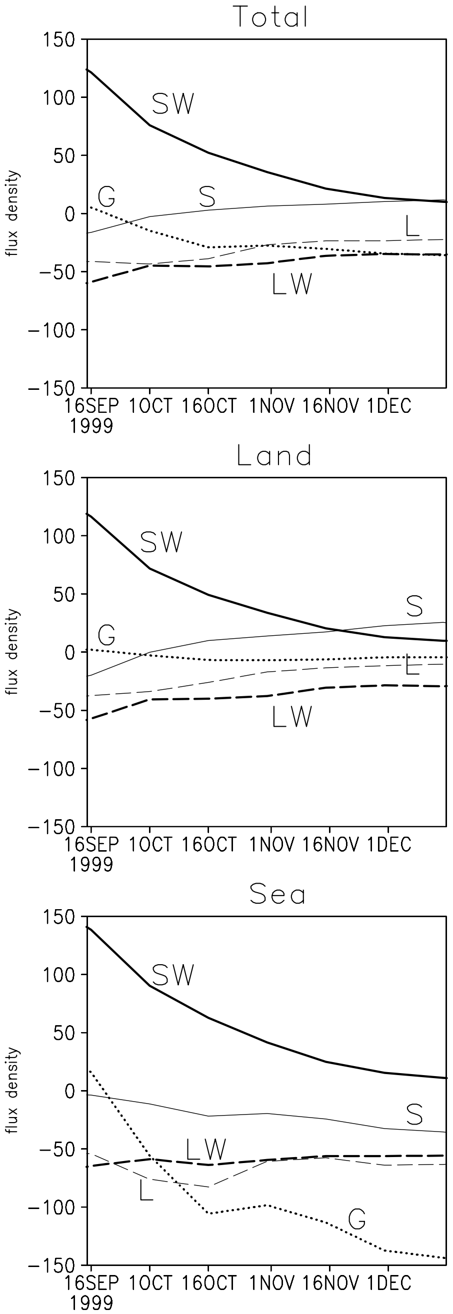

Fig. 3. Components of the mean surface heat balance over the catchment basin of the Baltic Sea; heating rates by shortwave radiation (SW), long wave radiation (LW), latent heat flux (*L*), and sensible heatflux (*S*). The net heat flux into the surface is denoted by (*G*). The graphs show 30-day averages in units of W m^{-2}.

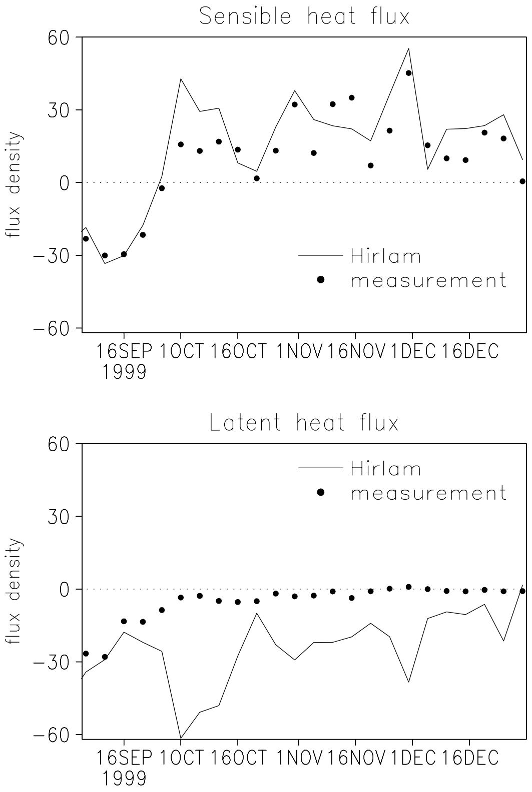

Fig. 4. Surface fluxes of sensible and latent heat at Hyytiälä, indicated by the cross in Fig. 1. The graphs show 5-day averages in units of W m^{-2}.

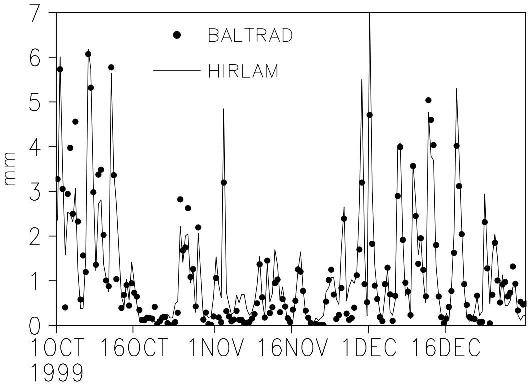

Fig. 5. Areal precipitation over the rectangular area shown in Fig. 1. The graphs show consecutive 12 hourly totals in mm. Solid lines refer to the present reanalysis. Dots indicate radar retrievals adjusted by rain gauges, provided by the Baltex Radar Data Centre.

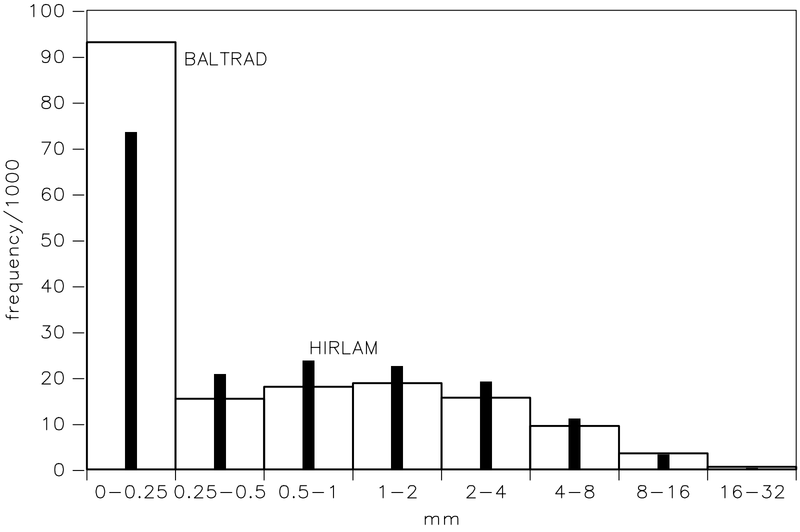

Fig. 6. Frequency distributions of twelve-hourly precipitation totals for all gridboxes within the rectangular area in Fig. 1 in October, November and December 1999. Thin black columns refer to the present reanalysis. Wide empty columns show radar retrievals adjusted by rain gauges, provided by the Baltex Radar Data Centre.

during summer (*G*) replaces SW as a heat source, and allows LW and *L* to remain nearly constant, while *S* even increases in magnitude, following the trend in the temperature contrast between air and sea.

In Fig. 4, model-generated surface fluxes of sensible and latent heat are compared to direct measurements at a SMEAR2 field station in Hyytiälä in southern Finland. The station, indicated by a cross in Fig. 1, belongs to the European-wide CARBOEUROFLUX network, and is described in Vesala *et al*. (1998). The fluxes are obtained from eddy covariance measurements above of a pine forest at a height of 23 m above ground. The model-fluxes represent a grid box with a mixture of 78% forest 14% open land and 8% lake. Bearing in mind these differences, a reasonable match is seen for the sensible heat flux. However, the latent heat flux is unrealistically strong in the model. Indeed, similar evaporation from winter-time land surfaces seems to be a general problem in the present model configuration, and probably contributes to the previously noted imbalance in the atmospheric water budget.

The verification of numerical precipitation forecasts in general is made difficult by the huge variability of precipitation in time and space. In general a large number of in situ measurements is needed to estimate the average precipitation over a model grid box. A network of radars provides virtually continuous observations, but obtaining accurate estimates of the precipitation at the surface using radars alone is very problematic. The BALTEX Radar Data Centre combines corrected rain gauge data with radar measurements over the catchment basin of the Baltic Sea. We use these data for verification of the predicted precipitation. Products and methodologies of the BALTEX Radar Data Centre (BALTRAD) are described in Michelson *et al*. (2000). The data used here consists of gridded consecutive 12-hourly precipitation sums with a horizontal resolution of 2 km. For the purpose of this study, the BALTRAD data are transformed by box-averaging to the HIRLAM-grid having a grid length of 22 km.

The entire radar network is shown in Fig. 1. The present study is restricted to the quadrilateral area covering southern Finland and surrounding seas, where the radar network is homogeneous, and the terrain rather flat. Fig. 5 shows the semi diurnal precipitation over this area. The correspondence between the two totally independent estimates is quite remarkable. Over these three months, the relative difference in accumulated precipitation is only about 10%, which is within the uncertainty of the BALTRAD retrieval.

Figure 6 shows frequency-histograms of semi diurnal precipitation during the three months for all (22 by 22 km) grid-boxes within the control area. HIRLAM is definitely able to reproduce the main features of the observed distribution. However, the model tends to underestimate the frequency of cases with no or very little precipitation. The number of weak and moderate cases is correspondingly overestimated.

Summary

We have presented the data assimilation system and a few results of the BALTEX regional reanalysis project. The assimilation system is essentially a modern limited area numerical weather prediction system, and could be used as such. Our results demonstrate that the system is capable of simulating the essential features of the energy and water cycles of the Baltic drainage basin. We find this encouraging, because the model has not been tuned to reproduce these cycles, but mainly to predict the atmospheric state. We therefore believe, that many of the physical processes involved in the climatic cycles of water and energy must be essentially correctly simulated by our system. We hope that this report may encourage others to make use of our data in their work, and to share their experiences with us.

Acknowledgements: We wish to express our thanks to Colin Jones and Björn Bringfelt for their invaluable advice and help in adopting modules from the regional climate model of the Rossby Centre, to Timo Vesala at the University of Helsinki (Department of Physical Sciences, Division of Atmospheric Sciences) for the flux measurements at Hyytiälä, and finally, to all our colleagues in the HIRLAM project for their help and encouragement. The BALTEX regional reanalysis is organized as a special project of the European Centre for Medium-Range Weather Forecasts and relies on the computing facilities and archives of that centre.

References

Bengtsson L. 1998. *Interim memorandum of understanding for the conduct of BRIDGE 1999–2001 in frame of BALTEX*. Int. BALTEX Secr., GKSS Forschungszentrum, Geesthacht, Germany, 58 pp.

Bringfelt B., Räisänen J., Gollvik S., Lindström G., Graham L.P. & Ullerstig A. 2001. The land surface treatment for the Rossby Centre Regional Atmosphere Climate Model — version 2. *SMHI Reports of Meteorology and Climatology* 98, Swedish Meteorological and Hydrological Intitute, 40 pp.

Browning K.A. & Gurney R.J. (eds.) 1999. *Global energy and water cycles*. Cambridge University Press, 292 pp.

Cuxart J., Bougeault P. & Redelsperger J.L. 2000. A turbulence scheme allowing for mesoscale and large-eddy simulations. *Q. J. R. Meteor. Soc*. 126: 1–30.

Fortelius C. 1995. Inferring the diabatic heat and moisture forcing of the atmosphere from assimilated data. *J. Clim.* 8: 224–239.

Gibson J.K., Kållberg P., Uppala S., Hernandez A., Nomura A. & Serrano E. 1999. *ERA-15 Description* (Version 2 January 1999), ECMWF Reading, UK, 84 pp.

Gustafsson N., Berre L., Hörnqvist S., Huang X.-Y., Lindskog M., Navascue B., Mogensen K.S. & Thorsteinsson S. 2001. Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus* 53A: 425–446.

Gustafsson N., Nyberg L. & Omstedt A. 1998. Coupling of a high-resolution atmospheric model and an ocean model for the Baltic Sea., *Mon. Wea. Rev.* 126: 2822–2846.

Haltiner G. & Williams R. 1980. *Numerical weather prediction and dynamic meteorology*, 2nd ed. John Wiley & Sons, New York, 447 pp.

Kalnay E., Kanamitsu M., Kistler R., Collins W., Deaven D., Gandin L., Iredell M., Saha S., White G., Woollen J., Zhu Y., Chelliah M., Ebisuzaki W., Higgins W., Janowiak J., Mo K.C., Ropelewski C., Wang J., Leetmaa A., Reynolds R., Jenne R. & Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* 77: 437–471.

Kain J.S. & Fritsch M.J. 1998. Multiscale convective overturning in mesoscale convective systems: Reconciling observations, simulations and theory. *Mon. Wea. Rev.* 126: 2254–2273.

Källen E. (ed.) 1996. *HIRLAM documentation manual System 2.5*, HIRLAM project, Swedish Meteorological and Hydrological Intitute, 120 pp.

Lindskog M., Gustafsson N., Mogensen K.S., Huang X.-Y., Yang X., Andræ U., Berre L., Thorsteinsson S. & Rantakokko J. 2001. Three-dimensional variational data assimilation for a limited area model. Part II: Observation handling and assimilation experiments. *Tellus* 53A: 447–468.

Ljungemyr P., Gustafsson N. & Omstedt A. 1996. Parameterization of Lake thermodynamics in a high resolution weather-forecasting model. *Tellus* 48A: 608–621.

Louis J.F., Tiedtke M. & Geleyn J.H. 1981. A short history of the PBL parameterization at ECMWF. In: *Workhops on Boundary Layer Parameterization*. ECMWF, Reading, U.K. pp. 59–80.

Michelson D.B., Andersson T., Koistinen J., Collier C.G., Riedl J., Szturc J., Gjertsen U., Nielsen A. & Overgaard S. 2000. BALTEX Radar Data Centre Products and their methodologies. *SMHI Reports Meteorology and Climatology* 90, Swedish Meteorological and Hydrological Intitute, 76 pp.

Rasch P.J. & Kristjansson J.E. 1998. A comparison of the CCM3 Model climate using diagnosed and predicted condensate parameterizations. *J. Climate* 11: 1587–1614.

Raschke E., Meywerk J., Warrach K., Andrae U., Bergström S., Beyrich F., Bosveld F., Bumke K., Fortelius C., Graham L.P., Gryning S.-E., Halldin S., Hasse L., Heikinheimo M., Isemer H.-J., Jacob D., Jauja I., Karlsson K.-G., Keevallik S., Koistinen J., van Lammeren A., Lass U., Launianen J., Lehmann A., Liljebladh B., Lobmeyr M., Matthäus W., Mengelkamp T., Michelson D.B., Napiórkowski J., Omstedt A., Piechura J., Rockel B., Rubel F., Ruprecht E., Smedman A.-S. & Stigebrandt A. 2001. The Baltic Sea Experiment (BALTEX): A European contribution to the investigation of the energy and water cycle over a large drainage basin. *Bull. Amer. Meteor. Soc.* 82: 2389–2413.

Savijärvi H. 1990. Fast radiation parameterization schemes for mesoscale and short-range forecast models. *J. Appl. Meteorol.* 29: 437–447.

Simmons A.J. & Gibson J.K. (eds.) 2000. *The ERA-40 project plan*, ECMWF Reading, UK, 60 pp.

Vesala T., Haataja J., Aalto P., Altimir N., Buzorius G., Garam E., Hämeri K., Ilvesniemi H., Jokinen V., Keronen P., Lahti T., Markkanen T., Mäkelä J.M., Nikinmaa E., Palmroth S., Palva L., Pohja T., Pumpanen J., Rannik U., Siivola Y., Ylitalo H., Hari P. & Kulmala M. 1998. Long-term field measurements of atmosphere-surface interactions in boreal forest combining forest ecology, micrometeorology, aerosol physics and atmospheric chemistry. *Trends in Heat, Mass & Momentum Transfer* 4: 17–35.

Wiin-Nielsen 1991: The birth of numerical weather prediction. *Tellus* 43AB: 36–52.

Yanai M., Esbensen S. & Chu J.-H. 1973. Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.* 30: 611–627.

Received 23 January 2002, accepted 8 May 2002