

DIAGNOSTIC SYSTEM OF ATMOSPHERE LOWER-LAYER FOR POLLUTION TRANSFER MODELLING

K. RUBINSTEIN, A. FROLOV, A. VAZHNIK, E. ASTACHOVA, I. ROSINKINA, D. KIKTEV and J. ALFEROV

Roshydrometcentre, Russia, 123342, Moscow, Bol. Predtechensky, 9-13. E-mail rusgmc@glas.apc.org

(Recibido septiembre 1995, aceptado febrero 1997)

Keywords: EMEP program, pollution transport, quasi 3-D multivariate analysis of meteorological data, precipitation objective diagnosis

ABSTRACT

A Diagnostic Atmospheric lower-layer System (SDA) has been elaborated in the Roshydrometcentre (Moscow) for meteorological data supply of pollution transfer modelling with the EMEP program. SDA allows for an operational receipt of dynamical characteristics (pressure, velocity of the wind, temperature on 1000, 925, 850, and 700 mb) and precipitation values in the points of a 150 km grid every 6 hours.

The dynamical SDA block is based on quasi-3D objective analysis, utilizing all available WWW observation data. Fields of hydrodynamic spectral forecast model at 0, 6, 12 and 18 hours are used as first guess. This model has a resolution of 4 levels up to 1.5 km and T40 spherical horizontal truncation.

All basic physical processes of the atmosphere, such as a cloud-radiation interaction, PBL parametrization, large-scale and convective precipitation, large-scale and convective clouds and processes on soil are included in the model.

The precipitation diagnosis is performed with the average in every EMEP grid point over neighbouring synoptic data in the 150 km radius circle. If no precipitation data are available the forecast of convective and large-scale precipitation by spectral model on 72 hours is used, giving the best precipitation forecast. The SDA system has become operational since January 1, 1994 with data storage.

An analysis of SDA run results is presented in this report. It is shown that the basic meteorological values are in satisfying agreement with one another and with empirical data. The wind velocity fields realistically reflect the large scale structure of the atmosphere circulation. The fields of the temperature and precipitation by synoptic analysis demonstrate a satisfactory localization of the main air mass and frontal zones.

RESUMEN

En el Roshydrometcentre de Moscú se ha elaborado un sistema diagnóstico de la atmósfera baja (SDA) para obtener datos meteorológicos con el objeto de modelar el transporte de contaminantes con el programa EMEP. El SDA permite la recepción operativa de características dinámicas (presión, velocidad del viento y temperatura a 1000, 925, 850 y 700 mb) y de los valores de precipitación en los puntos de una rejilla de 150 km cada 6 horas.

El bloque SDA se basa en un análisis objetivo cuasitridimensional utilizando los datos WWW de observación disponibles. Los campos del modelo de predicción espectral hidrodinámica a las 0, 6, 12 y 18 horas se utilizan como primera aproximación. Este modelo tiene una resolución de 4 niveles hasta 1.5 km y un corte esférico horizontal T40. En él se incluyen todos los procesos básicos de la atmósfera, tales como interacción nubes-radiación, parámetros PBL, precipitaciones convectivas y de gran escala, nubes convectivas y de gran escala y procesos en el suelo. El diagnóstico de precipitación se realiza dividiendo el promedio de cada punto en la rejilla EMEP entre los datos sinópticos vecinos en un círculo de 150 km de radio. Si no hay datos de precipitación convectiva y de gran escala el diagnóstico se hace con el modelo espectral para 72 horas, lo que resulta en el mejor pronóstico de precipitación. El sistema SDA con almacenamiento de datos ha estado en operación desde el 1 de enero de 1994.

En este trabajo se presenta un análisis de los resultados del SDA. Se demuestra que los datos meteorológicos básicos concuerdan unos con otros y con los datos empíricos. Los campos de velocidad del viento reflejan de manera real la estructura de gran escala de la circulación atmosférica. Por medio del análisis sinóptico los campos de temperatura y precipitación demuestran una localización satisfactoria de las principales masa de aire y zonas frontales.

INTRODUCTION

The intent to control the global ecological situation and, in particular, the already noticeable air pollution requires the precise description of the actual state of the lower troposphere.

This task cannot always be solved by traditional meteorological data processing. As it is known, the functioning chain "measurement-transmission-processing-final product-usage" nowadays is serving both the needs of weather forecast and identification of climate fluctuations.

One significant practical objective in the context of ecology is monitoring airborne pollution transport by atmospheric circulation. Since 1978, the Co-operative Programme for Monitoring and Evaluation of the long-range Transmission of Air pollutants in Europe (EMEP) has been implemented to meet this objective. Sulphur and nitrogen compounds are known to be transported at distance that might reach thousands of kilometers. Their dispersal and deposition depend on the circulation, stratification of the air masses and on the type of the underlying surface and its state.

The transboundary transport models used within the framework of the EMEP have the spatial resolution 150×150 km and a time step of 6 hours (Fig. 1A).

The Hydrometeorological Research Centre of the Russian Federation (Roshydromcentre) has no operational system for meteorological information processing with a mesh of such fine spatial resolution on the hemisphere scale. The data are processed by the Centre only for two synoptic observations per day (00 and 12 GMT). However, we wish to measure data four times a day. These conditions and the need to use all possible information suggested the need to develop a special system to diagnose the lower atmosphere state for the Meteorological Synthesising Centre-East (MSC-E).

Until 1984, MSC-E prepared the information on the atmospheric circulation using the procedure described in (Shapiro 1985). This procedure is based on synoptic and radiosonde air data and it gave good results for the European territory with a dense observation network.

Since the region of calculations has been expanded over the whole Northern Hemisphere excluding the tropical zone (Fig. 1B), it was impossible to use procedure Shapiro (1985) for the analysis of information for areas with few measurement data. Therefore, in 1993, a new System for Diagnosis of the lower Atmospheric state (SDA) was set up. This system uses all available meteorological data including satellite data; cat off interval in SDA is equal to 3 days. For more than a decade an automated information system of transboundary transport calculations (AISTTC) has been operating in MSC-E (Pressman 1985). Model pollutant fluxes in the atmosphere can be calculated through this system. The atmospheric boundary layer state and rain precipitation are the most important and variable input data for AISTTC.

According to procedure (Shapiro 1985), wind fields were calculated at 1000 and 850 gPa using geostrophical relation-

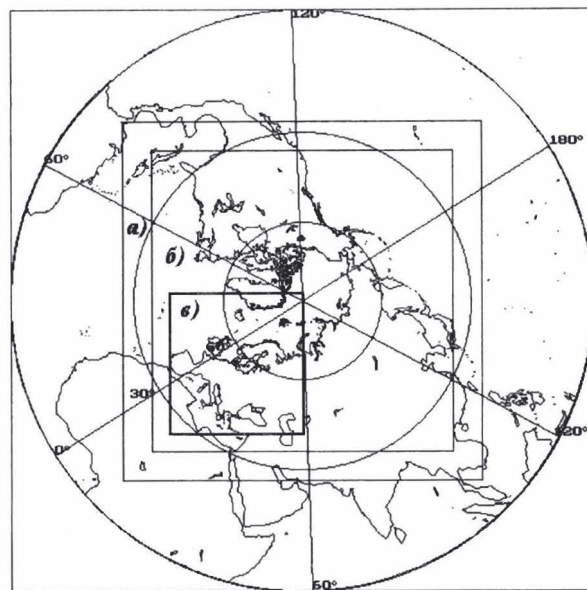


Fig. 1. The calculation of the air pollution transport model: A-EMEP region; B-expanded region

ships from geopotential fields reproduced on the basis of synoptic and upper air data. The geostrophic wind data obtained at grid nodes were corrected for wind speed measurements taken at the meteorological stations nearest to the nodes.

A short description of the meteorological data processing used in MSC-W (Meteorological - Syntes Centre, West, Oslo, Norway) is given in paper (Sanders 1993). As the initial information for transboundary air pollution transport model, this centre used both observations and results set by the numerical weather forecast model (NWF) of the Norwegian Meteorological Service.

In addition to data on the wind speed, surface layer temperature and precipitation, MSC-W used the information of the mixing layer height, friction stress, heat fluxes from the underlying surface and cloud amount (Table I). It should be noted that the accuracy of these atmospheric characteristics is doubtful. In this context we believe that the numerical experiments proposed for the evaluation of the transport model sensitivity of the listed parameters may be interesting.

A comparison of the meteorological information input shows that at equal spatial resolution in cases MSC-E, MSC-W previous system and MSC-W new system are considerably different, both in results obtained and methods employed.

CHARACTERISTICS OF OBSERVATIONAL DATA DELIVERED TO THE HIDROMETCENTRE VIA COMMUNICATION LINKS

Data obtained by various observational systems are the basis for the diagnosis of the current state of the lower atmosphere and the input to predict its variation. At present, via national and international communication, the Hydrometcentre

TABLE I. METEOROLOGICAL INFORMATION OBTAINED IN MSC-WEST(1993)

Physical parameters	Level	Time interval	Sources of information
Mean sea level pressure		6 hours	NWP* model
Horizontal components of wind speed	975 gHp	6 hours	NWP model
Mixing layer temperature	925 mbar	6 hours	NWP model
Relative air humidity	925 mbar	6 hours	NWP model
Air temperature	2 m above the absolute height of land-scope	6 hours	NWP model
Atmospheric precipitation	Underlying surface	6 hours total amount	NWP model, objective analysis of observation data (OA)
Cloud amount		6 hours	NWP model
Vertical component of wind speed	850 gPa	6 hours	NWP model
Mixing layer height		24 hours	OA
Friction stress	Underlying surface	6 hours	NWP model
Turbulent heat fluxes	Underlying surface	6 hours	NWP model
METEOROLOGICAL INFORMATION OBTAINED BY (MSC-EAST) (SHAPIRO 1985)			
Physical parameters	Level	Time interval	Sources of information
Horizontal components of wind speed	850 gPa6 1000 gPa	6 hours	OA (TEMP, SYNOP)
Atmospheric precipitation	Underlying surface	6 - hour total amount	Climatic values, OA
Air temperature	1000 gPa	6 hours	OA (SYNOP)
METEOROLOGICAL INFORMATION OBTAINED BY SDA (MSC-EAST, 1994)			
Physical parameters	Level	Time interval	Sources of information
Horizontal components of wind speed	850 gPa 925 gPa 1000 gPa	6 hours	OA (all kinds of observations), NWP model
Atmospheric precipitation	Underlying surface	6 - hour total amount	NWP model, OA
Air temperature	1000 gPa	6 hours	OA (all kinds of observations), NWP model

* NWP - Numerical Weather Prediction

receives measurement data at main (00; 12 GMT) and intermediate (6; 18 GMT) synoptic times. After preliminary control, messages decoding the data are recorded to the raw data base.

The most comprehensive and accurate data on the structure of a 30-km column of the atmosphere are provided through the upper air sounding (cod-'TEMP'). These data contain values of temperature, humidity, direction and speed of horizontal wind at 15 standard isobaric surfaces and their height above the sea level.

The stations located in the Northern Hemisphere send about 550 messages on average (one for each station) with 'TEMP' data for main synoptic times.

The 'TEMP' data over the continental areas are sufficiently dense, whereas only sparse data can be provided for oceanic areas since there are few island stations and few ships that carry out at the upper air sounding. Messages with TEMP

data for the intermediate periods are sent only from 15 stations located mainly in Europe.

Some stations carry out balloon measurements of wind speed and direction at standard isobaric surfaces (PILOT data). These stations are mainly located in Europe and America. The number of messages is approximately uniformly distributed with the main and intermediate synoptic periods; this number does not exceed 50-60 for one period.

The most complete information on meteorological values near the surface layer, is provided by synoptic stations (cod 'SYNOP'). The number of 'SYNOP' messages is about 3500, being approximately the same for main and intermediate observation time.

Considerably fewer messages are received from ship synoptic stations (SYNOPSIS SHIP data); 350-400 messages are received on average for one observation time.

Much information on atmospheric temperature and humidity characteristics is reported by polar-orbit meteorological satellites. On the basis of radiation intensity measurements in various spectral bands, centres of satellite data define the mean temperature of layer spacing between neighbouring standard levels, underlying surface temperature, and vertical profile of air humidity (SATEM data). Usually we have from 500 to 800 SATEM messages for each observation time. It is important for SDA that these observations are carried out over oceanic basins and in tropical regions, where only few upper air soundings are located. The total number of messages with SATOB data for one observation time varies from 500 to 1000.

The scheme of quasi-three-dimensional multivariate optimal interpolation developed for atmospheric pollution transport modelling presented in this paper processes data on pressure, temperature, wind and standard isobaric level heights (reports of SYNOP, SYNOP-SHIP, DRIBU, TEMP, PILOT, SATEM). The SATOB data are not used at present, though the lack of them can noticeably influence the interpolation accuracy in regions with few observations.

THE SCHEME OF DIAGNOSIS OF WIND, GEOPOTENTIAL AND TEMPERATURE FIELDS GIVEN BY OBSERVATIONS

At present many leading meteorological centres use schemes of meteorological data reproduction based on multivariate optimal interpolation (Lonnberg and Shaw 1983, Parrish D.F. and Derber J.C. 1991). Such schemes give the greatest possibility to take all peculiarities of measurement data into account and also permit the user to extrapolate the earlier observations due to the use of the numerical atmosphere model output as the first guess field (FGF). The latter improves significantly the reproduction accuracy.

At the beginning of the diagnosis scheme operation, data on pressure, temperature and humidity reported by SYNOP, SYNOP-SHIP and DRIBU are converted (by means of hydrostatics equation) into data of heights of the nearest standard isobaric levels. Wind measurements are also referred to these levels. Then, we calculate deviations of these data and TEMP, PILOT and SATEM data from FGF. The results are recorded into a special file sequentially in to the boxes with corresponding signs. For calculation of the deviations the FGF interpolation to stations with the fourth-order accuracy is used.

Data are processed box-by-box and level-by-level. First the 1000 gPa level is considered. For this level the interpolation is made not only to nodes of the grid but also to points where SATEM data are assumed. Hence, it is possible to pass over from the depths of slabs between isobaric levels of SATEM data to these levels heights and to use them in two-dimensional interpolation at other levels.

Let us have N observations of different types as a result of data selection at a certain level. We shall search the interpolation value A^a in form of the following linear combination of normalized deviations of observation data from the first guess (Frolov 1989, Vajnik and Frolov 1992):

$$\frac{A_i^a - A_i^p}{E_{Ai}^p} = \sum_{j=1}^N W_{ij} \frac{A_j^o - A_j^p}{E_{Aj}^p}, \quad (1)$$

where A implies any of such variables: Z (geopotential), U, V (wind speed components). Indices p, o indicate A belongs to the FGF or to observations; i - corresponds to the point where the interpolation is made to; j - corresponds to points, with observation data selected as influencing; E_{Ai} - statistical RMS (root-mean-square) of A element in point i .

The problem lies in the determination of weights W_{ij} .

The following notations are used:

$$\alpha_{Ai}^o = (A_i^o - A_i^t) / E_{Ai}^o; \alpha_{Ai}^p = (A_i^p - A_i^t) / E_{Ai}^p;$$

$$\alpha_{Ai}^a = (A_i^a - A_i^t) / E_{Ai}^a; \varepsilon_{Ai}^o = E_{Ai}^o / E_{Ai}^p; \varepsilon_{Ai}^a = E_{Ai}^a / E_{Ai}^p$$

A_i^t is the true value of A element in point i .

With allowance for these notations Eq (1) may be written as:

$$\alpha_{Ai}^a \varepsilon_{Ai}^a = \alpha_{Ai}^p + \sum_{j=1}^N W_{ij} (\alpha_{Aj}^o \varepsilon_{Aj}^o - \alpha_{Aj}^p) \quad (2)$$

Raising to the second power and using averaging over an ensemble of realizations, assuming non-correlation of α_{Ai}^p and α_{Aj}^o , we obtain:

$$\varepsilon_{Ai}^{a^2} = 1 - 2 \sum_{j=1}^N W_{ij} \langle \alpha_{Ai}^p, \alpha_{Aj}^p \rangle + \sum_{l=1}^N \sum_{j=1}^N W_{il} \left(\langle \alpha_{Ai}^p, \alpha_{Aj}^p \rangle + \varepsilon_{Al}^o \langle \alpha_{Aj}^o, \alpha_{Aj}^o \rangle \varepsilon_{Aj}^o \right) W_{ij} \quad (3)$$

where $\langle \alpha_{Ai}^p, \alpha_{Aj}^p \rangle$ means average over an ensemble of realization, in this case coinciding with the correlation of corresponding values. Write (3) as a matrix form:

$$\varepsilon_{Ai}^{a^2} = 1 - 2 \overline{W_i}^T \overline{P_i} + \overline{W_i}^T \overline{M W_i} \quad (4)$$

where forms of vectors $\overline{W_i}, \overline{P_i}$ and matrix $\overline{M}$ are easy to understand from (3). Now the problem lies in finding vector $\overline{W}$ minimizing (4). For this purpose we calculate

value $\frac{\partial \varepsilon_{Ai}^{a^2}}{\partial W_{ij}}$ and make it equal to zero. Thus, we obtain a set of

algebraic equations for the determination of optimal weights:

$$\overline{M W_i} = \overline{P_i} \quad (5)$$

or

$$\overline{W_i} = \overline{M}^{-1} \overline{P_i} \quad (6)$$

Using the obtained from (1) weights W_{ij} , we find A_i^a .

The final formula for the determination of $\varepsilon_{Ai}^{a^2}$ with allowance for (5) will be

$$\varepsilon_{Ai}^{a^2} = 1 - \overline{W_i}^T \overline{P_i}. \quad (7)$$

Thus, from assumed statistical characteristics of errors of the prediction and observations, it is possible to find the interpolated value in any given point and to evaluate the interpolation accuracy.

Suppose that the correlation over the horizontal of errors of isobaric surface height $\langle \alpha_{zi}^P, \alpha_{zj}^P \rangle$ depends only on the distance between points i and j , i.e. we take the hypothesis of uniformity and isotropy. Then we assume as in (Lonnberg and Shaw, 1983) that this correlation may be represented by a function of Gaussian type:

$$\langle \alpha_{zi}^P, \alpha_{zj}^P \rangle = \prod(r_{ij}) = \exp(-0.5 \frac{r_{ij}^2}{b^2}) \quad (8)$$

Where r_{ij} - distance between points i and j and b - selected parameter.

Following the ECMWF scheme, we take b to be locally constant but varying from box to box depending on the latitude ϕ :

$$b = \begin{cases} 600 \text{ km}, & 33^\circ \leq \phi \\ [900 - 10(\phi - 3^\circ)] \text{ km}, & 3^\circ < \phi < 33^\circ \\ 900 \text{ km}, & \phi \leq 3^\circ \end{cases}$$

If the climate is taken as the first guess field then $b = 1000$ everywhere.

For the calculation of the distance between points on a sphere in paper (Vaznik and Frolov 1992) a formula more simple than the one is suggested. Taking this formula into account expression (8) will be:

$$\prod(r_{ij}) = \exp\left(\frac{a^2}{b^2} \left[\sin\phi_i \cdot \sin\phi_j + \cos\phi_i \cos\phi_j \cdot \cos(\lambda_i - \lambda_j) - 1 \right]\right). \quad (9)$$

Statistical simulation of error of wind field prediction is a more complicated problem, because the connection of errors of the wind field and geopotential predictions is rather sophisticated. It is considered that geostrophical relationships between errors of wind field and geopotential predictions are fulfilled accurately enough. If U, V are the components of the horizontal wind with positive directions to the east and north, respectively, then we may write the following chain of equalities:

$$\begin{aligned} \langle \alpha_{zi}^P, \alpha_{zj}^P \rangle &= \frac{1}{E_{zi} E_{zj}} \langle U_i, U_j \rangle = \frac{g^2}{E_{zi} E_{zj} f_i f_j a^2} \left\langle -\frac{\partial z}{\partial \phi_i}, -\frac{\partial z}{\partial \phi_j} \right\rangle = \\ &= \frac{E_{zi} E_{zj} g^2}{E_{zi} E_{zj} f_i f_j a^2} \frac{\delta^2}{\delta \phi_i \delta \phi_j} \langle \alpha_{zi}^P, \alpha_{zj}^P \rangle = \frac{g^2}{b^2 f_i f_j} \frac{E_{zi} E_{zj}}{E_{zi} E_{zj}} \prod(r_{ij}) x \\ &= x \left(\frac{a^2}{b^2} \left[\sin\phi_i \cdot \cos\phi_j - \cos\phi_i \cdot \sin\phi_j \cdot \cos(\lambda_i - \lambda_j) \right] x \right. \\ &\quad \left. + \left[\cos\phi_i \cdot \sin\phi_j - \sin\phi_i \cdot \cos\phi_j \cdot \cos(\lambda_i - \lambda_j) \right] + \right. \\ &\quad \left. + \left[\cos\phi_i \cdot \sin\phi_j + \sin\phi_i \cdot \sin\phi_j \cdot \cos(\lambda_i - \lambda_j) \right] \right) \end{aligned} \quad (10)$$

where g is the gravitation acceleration, a - Earth radius, f_i - Coriolis parameter in point i , and E_{Ai} - MRS error of meteorological parameter A in point i .

Now assume that points i and j coincide. Then, on the one hand,

$$\langle \alpha_{U_i}^P, \alpha_{U_i}^P \rangle = 1,$$

On the other, from formula (10) we obtain: $\langle \alpha_{U_i}^P, \alpha_{U_i}^P \rangle = \frac{g^2}{b^2 f_i^2} \frac{E_{zi}^2}{E_{U_i}^2}$.

As a result we find that: $\frac{g}{b f_i} \frac{E_{zi}}{E_{U_i}} = 1$.

Hence we have:

$$\begin{aligned} \langle \alpha_{U_i}^P, \alpha_{U_i}^P \rangle &= \prod(r_{ij}) \cdot \left(\frac{a^2}{b^2} \left[\sin\phi_i \cdot \cos\phi_j - \cos\phi_i \cdot \sin\phi_j \cdot \cos(\lambda_i - \lambda_j) \right] x \right. \\ &\quad \left. + \left[\cos\phi_i \cdot \sin\phi_j - \sin\phi_i \cdot \cos\phi_j \cdot \cos(\lambda_i - \lambda_j) \right] + \right. \\ &\quad \left. + \left[\cos\phi_i \cdot \cos\phi_j + \sin\phi_i \cdot \sin\phi_j \cdot \cos(\lambda_i - \lambda_j) \right] \right). \end{aligned}$$

Similarly it is possible to obtain:

$$\begin{aligned} \langle \alpha_{V_i}^P, \alpha_{V_i}^P \rangle &= \prod(r_{ij}) \cdot \left[\cos(\lambda_i - \lambda_j) - \frac{a^2}{2} \cdot \cos\phi_i \cdot \cos\phi_j \cdot \sin^2(\lambda_i - \lambda_j) \right]; \\ \langle \alpha_{zi}^P, \alpha_{V_j}^P \rangle &= \frac{a}{b} \prod(r_{ij}) \cdot \left[\cos\phi_i \cdot \sin\phi_j \cdot \cos(\lambda_i - \lambda_j) - \sin\phi_i \cdot \cos\phi_j \right]; \\ \langle \alpha_{zi}^P, \alpha_{V_j}^P \rangle &= \frac{a}{b} \prod(r_{ij}) \cdot \cos\phi_i \cdot \sin(\lambda_i - \lambda_j); \\ \langle \alpha_{U_i}^P, \alpha_{V_j}^P \rangle &= \prod(r_{ij}) \cdot \sin(\lambda_i - \lambda_j) x \\ &\quad + x \left(\sin\phi_i - \frac{a^2}{b^2} \left[\cos\phi_i \cdot \sin\phi_j - \sin\phi_i \cdot \cos\phi_j \cdot \cos(\lambda_i - \lambda_j) \right] \cdot \cos\phi_j \right). \end{aligned}$$

Besides, the primary control of observation data contain much more accurate statistical control, if the number of influencing data is not less than three. The control algorithm is as follows:

Let data A_i^o be corrected. For this purpose we calculate $\frac{A_i^o - A_i^P}{E_{Ai}^P}$ by

means of optimal interpolation allowing for the elimination of A_i^o from the number of influencing data and the difference is

$$\text{determined: } \frac{A_i^o - A_i^a}{E_{Ai}^P} = \frac{A_i^o - A_i^P}{E_{Ai}^P} = \frac{A_i^a - A_i^P}{E_{Ai}^P}$$

Simultaneously, we calculate a square of statistical MRS error $\varepsilon_{Ai}^{oa^2}$:

$$\varepsilon_{Ai}^{oa^2} = \left\langle \frac{A_i^o - A_i^a}{E_{Ai}^p}, \frac{A_i^o - A_i^a}{E_{Ai}^p} \right\rangle = \left\langle \frac{A_i^o - A_i^t}{E_{Ai}^p} - \frac{A_i^a - A_i^t}{E_{Ai}^p}, \frac{A_i^o - A_i^t}{E_{Ai}^p} - \frac{A_i^a - A_i^t}{E_{Ai}^p} \right\rangle = \varepsilon_{Ai}^{o^2} + \varepsilon_{Ai}^{a^2}.$$

Here it is allowed for the fact that $\frac{A_i^a - A_i^t}{E_{Ai}^p}$ and $\frac{A_i^o - A_i^t}{E_{Ai}^p}$ do not correlate between each other, since value A_i^o is eliminated from the influencing data. If the equation is fulfilled

$$\left(\frac{A_i^o - A_i^a}{E_{Ai}^p} \right)^2 \geq C_{crit.} (\varepsilon_{Ai}^{oa^2} + 0.1), \quad C_{crit.} = 16,$$

A_i^o is eliminated.

In regions with dense network, it is necessary "to thin" observation data for assuring good motivation of the correlation matrix of the FGF matrix. The first "thinning" is made by the selection of influencing data allowing for their quality, representatives of all types of meteorological values and the best uniformity of these data in the vicinity. However, mechanical throwing away of other (not influencing) data results in the loss of useful information.

In the developed variant it is suggested to make "thinning" by the arrangement of superobservations with elevated accuracy.

According to this procedure superobservations are arranged in pairs for the same types of meteorological parameter.

Let observation data A_1^o, A_2^o with co-ordinates x_1, x_2 and MRS error of measurements E_1^o, E_2^o participate at the formation of superobservations A_s^s with co-ordinates x_s . Then assume:

$$\frac{A_s^s - A_s^a}{E_s^p} = W_1 \frac{A_1^o - A_1^p}{E_1^p} + W_2 \frac{A_2^o - A_2^p}{E_2^p}, \quad (11)$$

where A_k^p - FGF value of element A in point x_k with mean-root-square; E_k^p, W_i - weights.

If we introduce notations:

$$\alpha_s^s = \frac{A_s^s - A_s^t}{E_s^p}, \quad \varepsilon_s^s = \frac{E_s^s}{E_s^p},$$

where E_s^s - MRS of superobservation, then (11) may be rewritten as:

$$\alpha_s^s \varepsilon_s^s = \alpha_1^p + w_1 \alpha_1^o \varepsilon_1^s - w_1 \alpha_1^p + w_2 \alpha_2^o \varepsilon_2^s - w_2 \alpha_2^p. \quad (12)$$

Raise (12) to the second power and make averaging over the ensemble of realizations:

$$(\varepsilon_s^s)^2 = 1 + w_1^2 (\varepsilon_1^{o^2} + 1) + w_2^2 (\varepsilon_2^{o^2} + 1) - 2w_1 \langle \alpha_s^p \alpha_1^p \rangle - 2w_2 \langle \alpha_s^p \alpha_2^p \rangle + 2w_1 w_2 \langle \alpha_1^p \alpha_2^p \rangle. \quad (13)$$

It is required to find weights W_1 and W_2 in order that value $\varepsilon_s^{s^2}$ is minimum at $\langle \alpha_s^p \alpha_s^p \rangle = 1$. This condition is necessary for optimal interpolation of superobservation as a common one, which errors do not correlate with FGF errors.

Multiply (12) by α_s^p and statistically average:

$$w_1 \langle \alpha_s^p \alpha_1^p \rangle + w_2 \langle \alpha_s^p \alpha_2^p \rangle = 1. \quad (14)$$

Then minimum (13) under the condition of (14) is reached at:

$$w_1 = \frac{(\varepsilon_2^{o^2} + 1) \langle \alpha_s^p \alpha_1^p \rangle - \langle \alpha_s^p \alpha_2^p \rangle \langle \alpha_1^p \alpha_2^p \rangle}{(\varepsilon_2^{o^2} + 1) \langle \alpha_s^p \alpha_1^p \rangle^2 + (\varepsilon_1^{o^2} + 1) \langle \alpha_s^p \alpha_2^p \rangle^2 - 2 \langle \alpha_s^p \alpha_1^p \rangle \langle \alpha_s^p \alpha_2^p \rangle \langle \alpha_1^p \alpha_2^p \rangle}$$

$$w_2 = \frac{(\varepsilon_1^{o^2} + 1) \langle \alpha_s^p \alpha_2^p \rangle - \langle \alpha_s^p \alpha_1^p \rangle \langle \alpha_1^p \alpha_2^p \rangle}{(\varepsilon_2^{o^2} + 1) \langle \alpha_s^p \alpha_1^p \rangle^2 + (\varepsilon_1^{o^2} + 1) \langle \alpha_s^p \alpha_2^p \rangle^2 - 2 \langle \alpha_s^p \alpha_1^p \rangle \langle \alpha_s^p \alpha_2^p \rangle \langle \alpha_1^p \alpha_2^p \rangle} \quad (15)$$

and is equal to

$$\varepsilon_s^{s^2} = \frac{(\varepsilon_1^{o^2} + 1)(\varepsilon_2^{o^2} + 1) - \langle \alpha_1^p \alpha_2^p \rangle^2}{(\varepsilon_2^{o^2} + 1) \langle \alpha_s^p \alpha_1^p \rangle^2 + (\varepsilon_1^{o^2} + 1) \langle \alpha_s^p \alpha_2^p \rangle^2 - 2 \langle \alpha_s^p \alpha_1^p \rangle \langle \alpha_s^p \alpha_2^p \rangle \langle \alpha_1^p \alpha_2^p \rangle} \quad (16)$$

Since superobservations are formed out of sufficiently close observations, then it may be approximately taken:

$$\langle \alpha_s^p \alpha_1^p \rangle = \langle \alpha_s^p \alpha_2^p \rangle = \langle \alpha_1^p \alpha_2^p \rangle = 1.$$

Then from (15) and (16) we obtain:

$$w_1 = \frac{\varepsilon_2^{o^2}}{\varepsilon_1^{o^2} + \varepsilon_2^{o^2}}; \quad w_2 = \frac{\varepsilon_1^{o^2}}{\varepsilon_1^{o^2} + \varepsilon_2^{o^2}}; \quad \varepsilon_s^s = \frac{\varepsilon_1^{o^2} \varepsilon_2^{o^2}}{\varepsilon_1^{o^2} + \varepsilon_2^{o^2}}; \quad (17)$$

In order to find the co-ordinates of superobservations we link points x_1 and x_2 by the line of r length. Obviously, the superobservation should be on this line. Let the distance between x_1 and x_s be equal to r_1 , between x_2 and x_s equal to r_2 . Considering correlation functions for values of the same type and negligible value r , denominator ZN in expression (16) may be approximately written as

$$ZN = (\varepsilon_2^{o^2} + 1) \left(1 - \frac{r_1^2}{b^2} \right) + (\varepsilon_1^{o^2} + 1) \left(1 - \frac{r_1^2}{b^2} \right) + 2 \left(1 - \frac{r_1^2 + r_2^2 + r^2}{2b^2} \right)$$

It is clear that $\varepsilon_s^{s^2}$ defined by expression (16) will have the minimum in accordance with $r_1 (r = \text{const}, r_2 = r - r_1)$ if the denominator has the maximum (the numerator does not depend on r_1). Therefore settling derivative $\frac{\partial Z}{\partial r_1}$ equal to zero we obtain:

$$r_1 = \frac{r \varepsilon_2^{\sigma^2}}{\varepsilon_1^{\sigma^2} + \varepsilon_2^{\sigma^2}} \quad \text{or} \quad r_1 = W_2 r \quad (18)$$

similarly $r_2 = W_1 r$.

Thus, the superobservation can be calculated by formula (18) with weights

$$W_1 = \frac{E_2^{\sigma^2}}{E_1^{\sigma^2} + E_2^{\sigma^2}}; \quad W_2 = \frac{E_1^{\sigma^2}}{E_1^{\sigma^2} + E_2^{\sigma^2}}.$$

The superobservation co-ordinates allowing for (18) is determined in the following way: $x_s = W_1 \cdot x_1 + W_2 \cdot x_2$.

The MRS error of the superobservation "measurement" is taken to be equal to:

$$E^s = \frac{E_1^{\sigma^2} \cdot E_2^{\sigma^2}}{(E_1^{\sigma^2} + E_2^{\sigma^2})^{1/2}}$$

The algorithm of the presented procedure of the superobservation is realized sequentially in the following way: from the whole set of similar observations we take the pair most closely located to each other. This pair forms a superobservation if the following two conditions are met:

1. The distance between these data is less than 100 km;
2. These data values are such that:

k is determined empirically. In our case $k=4$.

A two-dimensional scheme of optimal interpolation compared to a three-dimensional one is not only less accurate but also provides less consistency of meteorological fields along the vertical. The three-dimensional scheme consumes 7-8 times more computer time than the two-dimensional one. Computer resources of the Roshydrometcentre do not allow us to use the three-dimensional scheme of optimal interpolation for operational and research activity.

This section deals with the procedure of bringing observations to neighbouring levels to eliminate a considerable extent of disadvantages of the two-dimensional optimal interpolation in comparison with the three-dimensional one practically without increase of computer time consumption.

THE PROGNOSTIC MODEL

To diagnose the current atmospheric state it is reasonable to use numerical modelling results of atmospheric processes based on observational data for the previous periods. The model will generate supplementary information on thermodynamic characteristics of the lower atmosphere which are not usually measured (wind speed vertical component, fluxes of sensible and latent heat, flux of momentum, characteristics of static stability, etc.).

In the Roshydrometcentre the operational numerical prediction of meteorological fields for 3-5 days is implemented by

an hemispherical atmospheric model (Kurbatkin *et al.* 1989, Frolov 1989). Hereafter, a brief description of the model is given.

The model is based on primitive equations of atmospheric hydrothermodynamics in spherical co-ordinates. Temperature (T), vortex (ξ) and divergence (D) of wind velocity and surface pressure (p_*) are used as prognostic variables.

The system is solved by a spectral numerical method when fields are presented as series of spherical harmonics. In the operational model version the truncation T40 is used.

In the model the atmospheric depth is divided into 15 layers up to the height of 10 gPa. Four of the main σ -levels are within the limits of the atmospheric boundary layer.

The time integration scheme is semi-implicit. The operational time step is 1800 s. In order to avoid the growth of fictitious high frequency oscillations in numerical solution we use a non-linear normal modes initialization procedure. This method is based on Machenhauer iteration scheme (Machenhauer 1977) modified in papers (Kurbatkin *et al.* 1988, Pakudov 1989).

The model includes detailed parametrizations of subgrid processes such as turbulence, cloud-radiation interaction, convection, large-scale condensation, processes in the surface soil layer. Let us consider the methods of these parametrizations.

The turbulent exchange of momentum, latent and sensible heat in the atmospheric surface layer is parametrized on the basis of Monin-Obukhov similarity theory with the usage of empirical functions of Businger. Sensible heat fluxes and momentum are written as in Louis (1979).

Horizontal diffusion is evaluated by a non-linear scheme of the fourth order which is modified by the introduction of enhanced viscosity for divergence and more efficient suppression of harmonics in the high-frequency part of the spectrum in the stratosphere.

Cloud-radiation interaction is parameterized in the model by the scheme given in (Geleyn 1979, Astachova 1992). Two stream assumptions and Eddington's approximation are used for the transition to the multilayer atmosphere. For the determination of radiation fluxes the map path length method is used.

It is supposed that cloudiness may occur in any atmospheric layer but in the surface boundary layer. Cloud amount C is calculated by a diagnostic formula depending on relative R and critical humidity R_{cr} . The critical humidity depends only on value $\sigma = p/p_*$.

Radiation fluxes are calculated separately for cloudy and cloudless parts of the atmosphere. Total fluxes are presented as their weighted sum.

Moist convection is considered in the model with the help of the Kuo-scheme (Kuo 1965, 1974, Anthes 1974).

The model contemplates the possibility of rain evaporation according to Kessler's scheme (Kessler 1969). It is assumed that evaporation depends on saturation deficit and density of the rain. The occurrence of rainfall or of snow fall is determined according to surface temperature. The large-scale condensation is considered in the model in a usual way.

CALCULATION OF TOTAL PRECIPITATION IN A GRID CELL

The precipitation amount in grid cells is calculated in two stages. The first one includes the acquisition and accumulation of precipitation data at stations. The second one implies calculation of total precipitation in grid knots on the Earth surface on the basis of obtained data.

Precipitation data are taken from synoptic cables SYNOP. Some information can be also provided by cables SYNOP-SHIP. As a rule, ships do not make reports on precipitation because they are moving and it is difficult to make reference of the geographical location and the time period t . Sometimes stand-still weather vessels, however, provide these data.

Reports SYNOP provide sufficient data for Europe, middle and southern parts of Siberia. Less data are available for North America, Mongolia and China. There are sparse data for northern and high mountain parts of Asia and North Africa. Data for oceans are practically non-existent.

Precipitation data reported by synoptic messages present the same precipitation for a certain period of time. The period duration is also coded and it may change from one to 24 hours. Therefore, an algorithm for data distribution with synoptic periods during 24 hours must be developed. In this paper, precipitation data collected 3 hour before and 3 hours after precipitation are ascribed to a synoptic period.

We used the following algorithm of the received data distribution with the indicated time intervals. For each message with precipitation data we determined synoptic observation times covered by the data collection period. Then, for each moment of time we found out whether any previous message with relevant data ascribed to any of 6 hours observations. If so, the earlier received precipitation sum for "occupied" hours was subtracted from the latest data. After that, the remained sum was distributed with "unoccupied" hours. It was assumed that during the period of data collection the precipitation fell uniformly in time and with constant intensity. Using this algorithm, we succeeded in processing duplicated reports from the same geographical locations (ground synoptic stations and weather vessels).

For each observation time, the data collection algorithm provides about 4000-4300 reports on precipitation from ground stations and 3-4 (more seldom 7-9) ship reports for the whole Northern Hemisphere.

While making objective analysis of meteorological parameter by means of an appropriate interpolation technique, we found their values in grid knots. In view of the fact that a precipitation sum is a discontinuous spatial function and the data are not instantaneous, a usual technique of objective analysis of meteorological fields cannot be applied.

For the MSC-E long-range transport modelling, we assigned to a regular grid node mean arithmetic value of sums of precipitation recorded at stations in a square with centre in a given knot with a side equal to the distance between knots.

As it was said above there are vast areas with sparse observations. About 4000 out of 6889 grid knots remain undeter-

mined after the averaging over squares procedure. Therefore, in order to fill the data gaps, it is necessary to use some supplementary information.

To re-establish continuous precipitation fields over the whole hemisphere by observation data, the background field (FGF) may be used as it is done in the procedure of objective analysis. In modern data assimilation systems as a rule for this field, the earlier calculated hydrodynamic 6-12 hours forecast is taken. The hydrodynamic precipitation forecasts gain appropriate quality for users only after more than 24-hour integration of the model and only for advanced ones. Therefore the usage of 6-hour hydrodynamic forecast as FGF for re-establishment of precipitation fields is not optimal (since the forecast quality of this characteristic does not still achieve an acceptable level). The application of climatic norms is accompanied by great inaccuracies in subsequent calculations.

Judging by the experience gained, the "best quality" of hydrodynamic forecast of precipitation is achieved when it is made 48-60 hours in advance using models of this kind (Roads *et al.* 1993). An analysis of daily precipitation forecast for the central european part of Russia indicated that the model provides sufficiently good results for regions with round-the clock precipitation. A comparison with measurements made in a great number of meteorological stations located within several cells of the atmospheric models (moderate latitudes, summer time), showed that in 75% of the cases with no precipitation, the model gave zero and in about 65% of the cases the prediction gave 1 mm/day when it rained over the major part of a square.

In this context, it was decided to develop a simplified precipitation system analyzing the combined hydrodynamic forecast 48, 54, 60, 66 hours in advance and the measurement data.

ANALYSIS OF RESULTS

SDA operational results provide meteorological fields which are the input for transboundary pollution transport model AISTTC. Since November 1994, the required fields have been accumulated on magnetic tapes. Some examples (patterns of the fields built up by SDA procedures) are considered in this section. Besides, some meteorological fields obtained by the methods used in MSC-E before 1994 are compared with those obtained by SDA technology.

The input data for the transboundary pollution transport model AISTTC are the wind fields at 1000, 925, 850 gPa levels, air temperature at 1000 gPa and precipitation. Let us consider a number of charts for 12 GMT, October 31, 1993 derived from SDA.

Fig. 2 (a,b,c) gives information on the geopotential surface field heights, as well as an actual location (taken from synoptic charts) of cyclone and anticyclone centres at the sea level. It is evident that the fields reflect the large-scale atmosphere circulation structure. Secondly, the location of the main baric systems are close to reality. Fields of the geopotential heights are not used in the input for the MSC-E model, therefore, they are

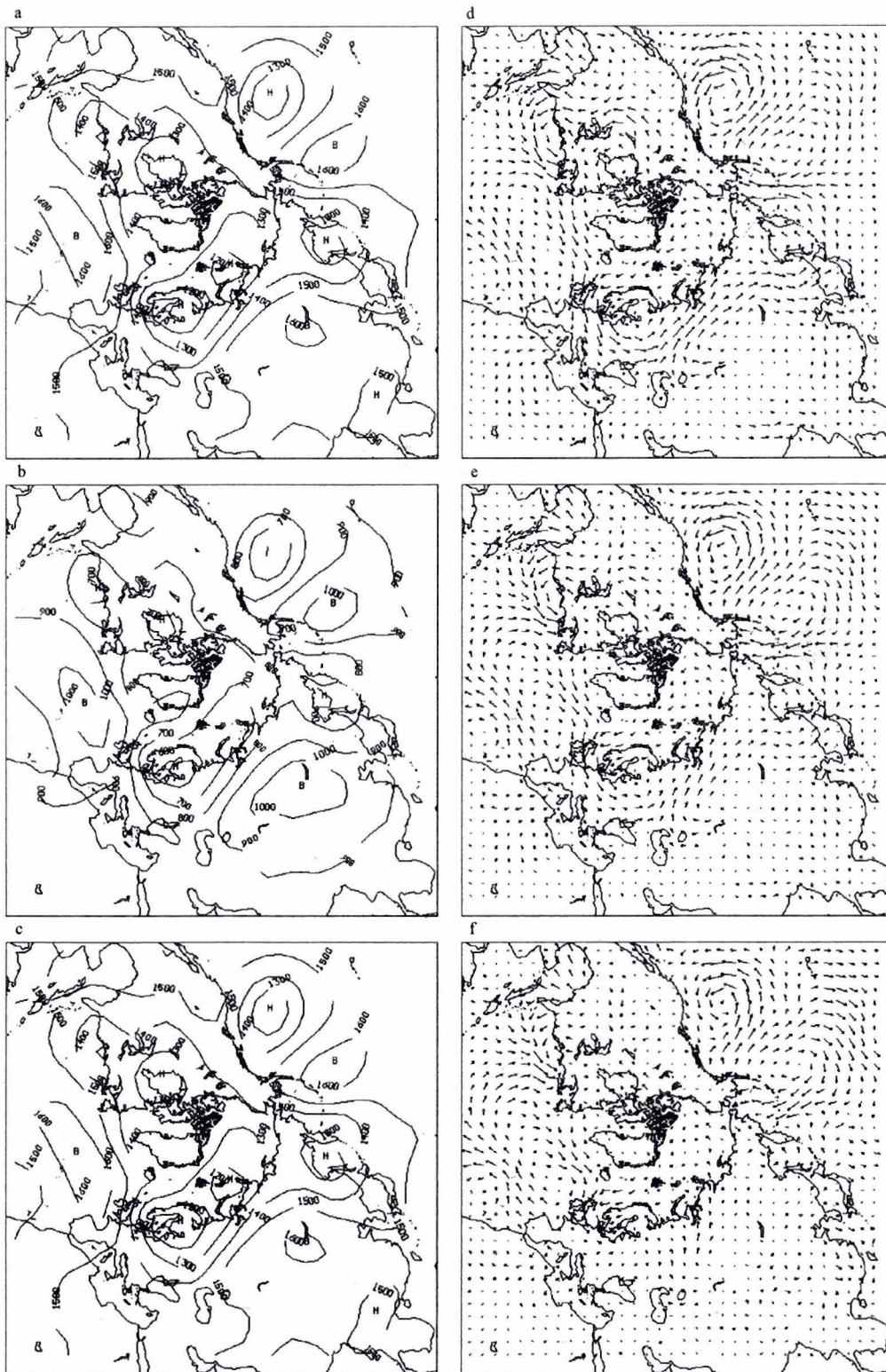


Fig. 2. Diagnosis of geopotential and wind fields for 12 GMT, 31.10.93:

a - field of 850 gPa geopotential surface height;
 b - field of 925 gPa geopotential surface height;
 c - field of 1000 gPa geopotential surface height;

d - wind field on 850 gPa geopotential surface;
 e - wind field on 925 gPa geopotential surface;
 f - wind field on 1000 gPa geopotential surface

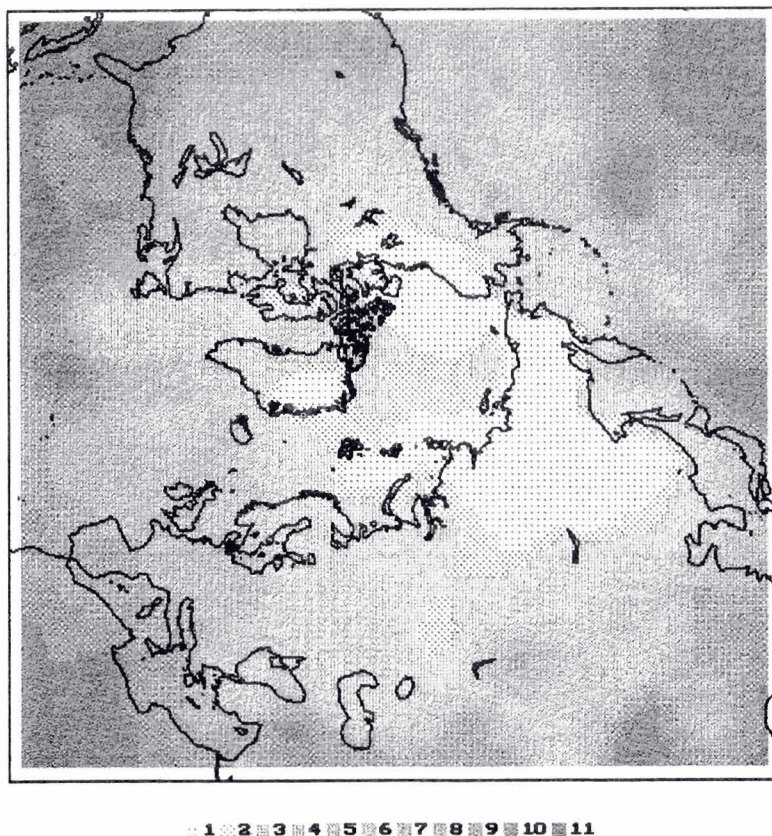


Fig. 3. Diagnosis of temperature field on H1000 for 12 GMT 31.10.93

not kept in archives in SDA. The analysis of these figures is important however, since the geopotential restoration is the base for the reproduction of the total set of meteorological characteristics at the appropriate levels.

Fields of the horizontal wind speed vector are shown in Fig. 2 (d,e,f). Obviously they not only fully correspond to the circulation peculiarities of the baric systems, but also reflect the friction effect which increases as the underlying surface is approached. The later is expressed by the decrease and turn of the wind speed vector at the lowest level.

The 1000 gPa surface level has a number of peculiarities. First, in many regions (corresponding to cyclones with pressure less than 1000 gPa at the sea level) values of 1000 gPa height become negative (Fig. 2c). Second, over vast areas with boundaries and location depending to a great extent on baric system locations, and practically all over the mountain regions, this surface appeared to be below the Earth surface.

Temperature fields at 1000 gPa level (Fig. 3) are presented in gradation (the first gradation corresponds to values below -12 degree, the second between -12:-8; the third between -8:-4; the fourth between -4:0; the fifth between -0:+4; the sixth between +4:+8; the seventh between +8:+12; the eighth between +12:+16; the ninth between +16:+20; the tenth between +20:+24; eleventh - above 24°) background.

It is clear that the fields correspond to a large-scale thermal structure of baric systems. One can see regions of warm cy-

clones sectors and cold regions at their rear (for example, for cyclones in the area of the Aleutian Islands, over the Norwegian Sea, North Atlantic and Western Siberia). The temperature increase detected in the background along the cyclones tracks of south-western direction also represents the reality.

Figure 4 shows the analysis of precipitation for 00 GMT of January 11, 1994. This figure illustrates the quality of the hydrodynamic forecast of precipitation. It also shows the position of baric systems and atmospheric fronts. It is noticeable that the location of the most intensive precipitation regions coincides with the main atmospheric fronts and the intensive baroclinic instability zones. The main drawback is the forecast of some weakly precipitation in the regions of powerful anticyclones (Yakutia, Central Asia), while actually only fogs have been observed in those regions. Nevertheless this forecast was estimated as a high standard prognosis. The result is a combination of prognostic and actual precipitation fields.

Thus, meteorological fields, being the SDA output, are of a realistic character as well as consistent and have no apparent errors in regions with no observations. The final conclusion of SDA output quality can be made when statistically significant data series are accumulated. It is also important to make a number of numerical experiment transboundary transport models using meteorological fields produced by SDA.

We assume that the system configuration did not change

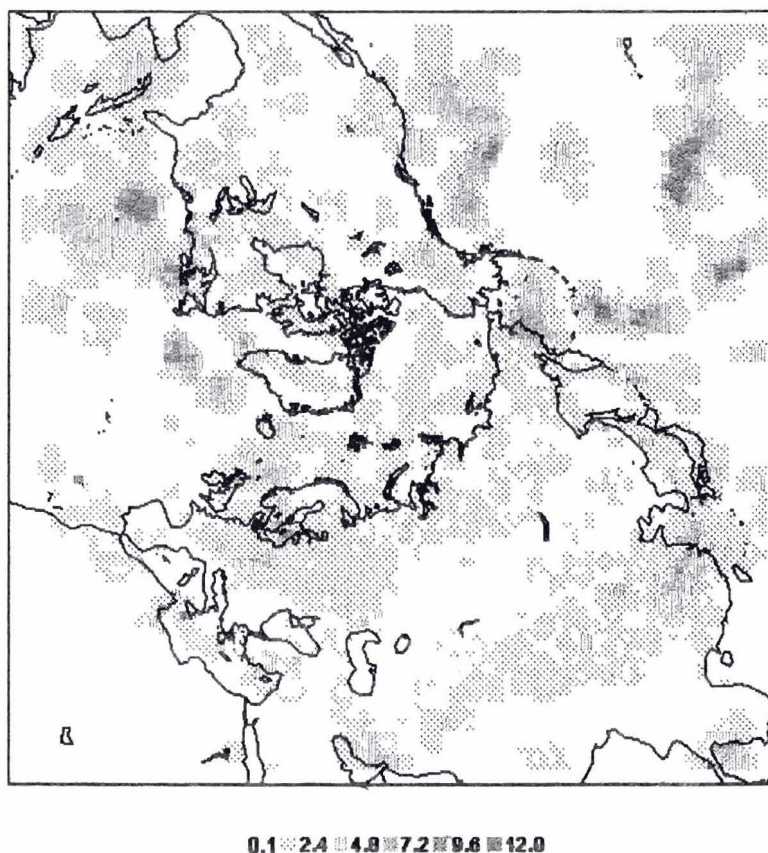


Fig. 4. The result of the superposition of prognostic and actual precipitation sums for 9-15 GMT, December 18, 1993

since 1994. Now a new version of the system is being developed. The main directions of the development will be directed to the optimization of the operational technology and to broaden the list of parameters to be analysed.

The spatial grid resolution has remained the same, but the calculation territory will be expanded to 99×99 grid points instead of 83×83 , and will cover almost the whole Northern Hemisphere.

REFERENCES

- Anthes R.A. (1977). A cumulus parametrization scheme utilizing a one-dimensional cloud model. *Mon. Wea. Rev.* 105, 270-286.
- Astakhova E.D. (1992). On approximate of diurnal cycle of radiation in general circulation atmospheric model. *Meteorol. Hydrol.* 9, 23-26
- Frolov A.V. (1989). On operational hydrodynamic medium range forecast in the USSR Hydrometcentre. *Proc. of the International Symposium on Methods of Medium Weather Forecast*, L., Hydrometeoizdat, pp. 134-150.
- Geleyn J.P. and Hollingsworth A. (1979). An economical analytical method for the computation of the interaction between scattering and line absorption of radiation. *Contr. Atmos. Phys.* 52, 1-16.
- Kessler E. (1969). On distribution and continuity of water substance in atmospheric circulation. *Met. Monogr. Amer. Met. Soc.* 10, 3-84.
- Kuo H.L. (1965). On formation and intensification of tropical cyclones through latent heat release by cumulus convection. *J. Atmos. Sci.* 22, 40-63.
- Kuo H.L. (1974). Further studies on the parametrization of the influence of cumulus convection on large-scale flow. *J. Atmos. Sci.* 31, 1231-1240.
- Kurbatkin G.P., Astakhova E.D., Krupchatnikov V.N., Rabinin V.E., Salnik V.A., Smirnov V.D. and Frolov A. V. (1987). A model of medium-range weather forecast. *Transactions of the USSR Academy of Sciences*, v.194, iss.2.
- Kurbatkin G.P., Zulunov S.M., Pokudov A.V. and Fracov A.V. (1988). Non-linear adjustment of wind and pressure fields during the numerical weather forecast using primitive equations. *Meteor. Hydrol.* 12, 5-13.
- Lonnberg P. and Shaw D. (1983). ECMWF data assimilation. *Scientific Documentation. ECMWF Met.Bull.*, M1.5/1.
- Louis J.F. (1979). A parametric model of vertical add fluxes in the atmosphere. *Bound. Layer Met.* 2, 187-202.

- Machenhauer B. (1977). The dynamic of gravity oscillations in a shallow water model with application to normal model initialization. *Contr. Atmos. Phys.* 50, 253-271.
- Parrish D.F. and Derber J.C. (1991). The National Meteorological Centre's spectral statistical interpolation analysis system. Office Note 379, U.S. Department of Commerce, NOAA, NMS, 42 p.
- Pokudov A. V. (1989). Relaxation method of non-linear initialization using normal modes. *Meteorol. Hydrol.* 12, 109-111.
- Pressman A. Ja. and Bronin B.N. (1985). The first turn of an automatized information system for transboundary air pollution transport. *Proc. of the Institute of Applied Geophysics* 62, 3-7.
- Roads J.O., Maisel T.N. and Alpert J. (1993). Further evaluation of National Meteorological Center's Medium-Range Forecast Model Precipitation Forecasts. *Wea. Forecast.* 6, 483-495.
- Sandnes H. (1993). Calculated budgets for airborne acidifying components in Europe, 1985-1992. EMEP/MSC-W Report 1/93, July.
- Shapiro M. Ja. (1985). Objective analysis of the wind field on the basis of data on geopotential and wind. *Proc. of the Institute of Applied Geophysics* 62, 59-64.
- Vazhnik A.I. and Frolov A. V. (1992). A discrete system of four-dimensional assimilation of data atmosphere state over the Northern Hemisphere. *Meteorol. Hydrol.* 11, 37-48.