Intercomparison study of atmospheric mercury models: 1. Comparison of models with short-term measurements

Alexey Ryaboshapko [a], O. Russell Bullock Jr. [b,d], Jesper Christensen [c], Mark Cohen [d], Ashu Dastoor [e], Ilia Ilyin [a,*], Gerhard Petersen [f], Dimiter Syrakov [g], Richard S. Artz [d], Didier Davignon [e], Roland R. Draxler [d], John Munthe [h]

[a] *Meteorological Synthesizing Center — East of EMEP, Leningradsky Pr., 16-2, Moscow 125040, Russia*
[b] *U.S. EPA Office of Research and Development, Research Triangle Park NC 27711, USA*
[c] *Department of Atmospheric Environment, National Environmental Research Institute, PO Box 358, Roskilde, Denmark*
[d] *NOAA Air Resources Laboratory, 1315 East West Highway, Silver Spring MD 20910, USA*
[e] *Air Quality Research Branch, Meteorological Service of Canada, Environment Canada, Dorval, Quebec, Canada*
[f] *GKSS — Research Centre, Max-Plank-Strasse 1, D-21502 Geesthacht, Germany*
[g] *National Institute of Meteorology and Hydrology, Tzarigradsko chaussee 66, 1785 Sofia, Bulgaria*
[h] *Swedish Environmental Research Institute, Dagiamningsgatan 1, PO Box 47086, S-40758 Göteborg, Sweden*

Abstract

Five regional scale models with a horizontal domain covering the European continent and its surrounding seas, one hemispheric and one global scale model participated in an atmospheric mercury modelling intercomparison study. Model-predicted concentrations in ambient air were compared against mercury species observed at four monitoring stations in Central and Northern Europe and a station on the Irish west coast. The modelled concentrations of total particulate mercury (TPM) were generally consistent with the measurements at all sites. The models exhibited significant ability to simulate concentrations of gaseous elemental mercury (GEM), but some of the short-duration peaks at the Central European stations could not be consistently reproduced. Possible reasons for these discrepancies include (1) errors in the anthropogenic emissions inventory utilized; (2) coarse spatial resolution of the models; and (3) uncertainty of natural and re-emitted mercury sources. The largest discrepancies between measurements and modelled concentrations were found for reactive gaseous mercury (RGM). For these models, the uncertainty in predicting short-term (two-week episode) variations of mercury species in air can be characterized by the following overall statistics: 90% of the results for TGM are within a factor of 1.35 of the measurements; for TPM, 90% are within a factor of 2.5; and for RGM, 90% are within a factor of 10.

1. Introduction

The Cooperative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe (EMEP) was established by the

* Corresponding author. Tel./fax: +7 495 614 45 94.
  E-mail address: ilia.ilyin@msceast.org (I. Ilyin).

1979 Convention on Long-Range Transboundary Air Pollution on Heavy Metals (signed currently by 36 countries). One of the primary objectives of EMEP is to provide (using appropriate models) estimates of transboundary fluxes and deposition of heavy metals within the geographical scope of EMEP. Hence, the modelling results are intended to be a basis for the countries to make decisions that are potentially very expensive. The Meteorological Synthesizing Centre East (MSC-E) of EMEP has a responsibility to perform these model calculations. A natural question arises — to what extent can the countries trust the modelling results? The existing mercury monitoring networks are too scarce and do not cover the modelling domain. Moreover, some essential parameters of mercury transport and deposition are not measured by current monitoring systems at all (short-lived compounds, dry deposition etc.). Hence, one useful activity that can be done to begin to establish trust is to compare the models among themselves to get a measure of the overall modelling uncertainty. In addition, the model intercomparisons are essential prerequisites for the development and application of operational models.

Many numerical models of different types have been developed to evaluate the atmospheric transport and deposition of heavy metals on local, regional and global scales. They are widely used as purely scientific instruments or as applied methods to solve environmental problems of local or national scopes. It is understandable that the models can differ from each other depending on their complexities and their tasks. Intercomparison studies have already been completed for lead (Sofiev et al., 1996) and cadmium (Gusev et al., 2000). This mercury model intercomparison study was initiated in 1999 to be carried out in three stages: (I) comparison of algorithms for the physico-chemical transformations of mercury species in a cloud/fog environment (Ryaboshapko et al., 2002); (II) comparison of model results with observations of mercury concentrations in air during two short-term episodes; and (III) comparison of model results with observed monthly and annual means of mercury concentrations in air and precipitation as well as comparison of estimated deposition fluxes over the entire EMEP domain and over selected European countries.

Here we report the results of stage II (the results of stage III are presented in a companion paper, Ryaboshapko et al., 2007). The main task of this stage was to characterize the uncertainty in modelling short-term (days–weeks) mercury concentration variations in the atmosphere. Three mercury physico-chemical forms were considered: gaseous elemental mercury (GEM),

total particulate mercury (TPM), and reactive gaseous mercury (RGM). Participating models performed simulations of the fate and transport of atmospheric mercury emitted from a common emissions inventory for two different multi-week episodes (one in 1995 and one in 1999). Model predictions were compared against measurements at five regional monitoring sites in Europe for these two episodes. It should be noted that in some cases the measurements give concentration values for total gaseous mercury (TGM = GEM + RGM). Because the RGM contribution to TGM is very small, below we consider the corresponding measurement results as GEM. Only a brief summary of the methodologies and results can be given here, but more extensive details are available elsewhere (Ryaboshapko et al., 2003).

2. Participating models

A brief summary of the main characteristics of the participating models is given in Table 1, along with references for more detailed descriptions. The participating models include a regional Lagrangian formulation and Eulerian approaches on regional, hemispheric and global scales. All the models employ extensive gas- and aqueous-phase chemical mechanisms and explicitly track numerous species concentrations. In these algorithms, mercury species are transferred among different phases in the atmosphere (e.g., gas, aqueous, particles in air, and particles in liquid) according to equilibrium and/or mass transfer considerations. The models simulate the chemistry of mercury within the various phases in the atmosphere, including the complexation of mercury species with other compounds in the aqueous phase. Some of these reactions oxidize elemental to ionic mercury and some reduce ionic to elemental mercury. As deposition processes can be dramatically different for different mercury forms, an accurate description of these transformation reactions is vitally important to predict the fate of mercury species in the atmosphere. The chemical schemes employed by the various models are very similar, largely consistent with the schemes described in Ryaboshapko et al. (2002).

Modelling of mercury chemical transformations in the atmosphere requires the knowledge of concentrations of several reactants and of a number of geophysical parameters (Ryaboshapko et al., 2002). Table 2 describes the chemical parameters and/or data sets that were used as the inputs in this intercomparison study. Note that some of the models used the pre-simulated sulphur dioxide and ozone concentrations described in

Table 1
Characteristics of participating models

Model name	ADOM	CMAQ	DEHM	EMAP	GRAHM	HYSPLIT	MSCE-HM
Model type	Eulerian	Eulerian	Eulerian	Eulerian	Eulerian	Lagrangian	Eulerian
Model domain	Europe	Central and Northern Europe	Northern Hemisphere	EMEP domain[a]	Global	Northern Hemisphere	EMEP domain[a]
Horizontal resolution (km, unless noted differently)	55×55	36×36	50×50 150×150[c]	50×50	$1° \times 1°$	$2.5° \times 2.5°$[b]	50×50
Model top height (km)	10	15	15	5	30	25	3.9
Source of meteorological data	HIRLAM, off-line	ECMWF TOGA reanalysis, MM5, off-line	ECMWF TOGA reanalysis, MM5, off-line	NCEP/ NCAR reanalysis, SDA, off-line	Canadian Meteorological Centre, on-line	NCEP/ NCAR reanalysis, off-line	NCEP/ NCAR reanalysis, SDA off-line
Boundary conditions[d]							
GEM (ng m^{-3})	1.5	1.7	1.5	1.5		1.5[e]	1.6–1.7[f]
TPM (pg m^{-3})	20	17	0	10	None	10[e]	20
RGM (pg m^{-3})	2	17	0	10		5[e]	0
Dry deposition[g]							
GEM	–	–	–	–	DDV	–	DDV
TPM	RA	RA	RA	DDV	RA	RA	RA
RGM	RA	RA	RA	DDV	RA	RA	DDV
Wet deposition[i]	CM	SC	SC	SC	CM	SC	SC
Gas-phase oxidation agents	O_3	O_3, H_2O_2, Cl_2, OH	O_3	O_3, OH	O_3	O_3, H_2O_2, Cl_2, HCl, OH	O_3[j]
Aqueous-phase oxidation agents	O_3	O_3, OH, HOCl, OCl$^-$	O_3	O_3	O_3	O_3, OH, HOCl, OCl$^-$	O_3
Aqueous-phase reduction agents	$SO_3^=$	$SO_3^=$, hv, HO_2	$SO_3^=$	$SO_3^=$	$SO_3^=$	$SO_3^=$	$SO_3^=$, HO_2
References	Petersen et al. (2001)	Bullock and Brehme (2002)	Christensen et al. (2004)	Syrakov (1995)	Dastoor and Larocque (2004)	Cohen et al. (2004)	Ilyin et al. (2002)

[a]EMEP domain includes Europe, partly the North Atlantic and the Arctic oceans, Northern Africa, and part of Middle East (www.emep.int); [b]resolution of meteorological data; [c]for EMEP domain and outside the domain, respectively; [d]presented values are for the ground level, normal conditions; [e]applied throughout the model domain; [f]depending on a boundary; [g]RA is a resistance approach, DDV is a simple dry deposition velocity approach; [h]empirical function of surface roughness and friction velocity; [i]CM is cloud microphysics, SC is a scavenging coefficient approach; [j]temperature dependent reaction rate.

this table, and other models included photochemical simulations (e.g., to derive estimates for ozone concentrations) and/or sulphur fate and transport simulations (to derive estimates for SO₂ concentrations).

It is well known that global sources can affect the mercury concentrations at any given location. By definition, models with non-global model domains cannot explicitly account for such influences. Accordingly, as is the standard practice, the regional and hemispheric Eulerian models (CMAQ, ADOM, EMAP, MSCE-Hg) utilized the various boundary conditions shown in Table 1 in an attempt to account for sources outside their model domains. The formulation of regional Lagrangian models does not generally allow such boundary conditions. For these types of models, a constant background concentration is typically added

throughout the model domain (e.g., Shannon and Voldner, 1995; Petersen et al., 1995; Bullock et al., 1998; Forlano et al., 2000). As shown in Table 1, the regional Lagrangian model involved in this exercise (HYSPLIT) used this same procedure to account for the influence of sources outside its domain. The use of a background concentration for GEM and TPM is consistent with their well-known ability to be transported over long distances in the atmosphere. The use of a background for RGM is supported by measurements for RGM in Europe (e.g., Munthe et al., 2003). We note that the background concentrations utilized for RGM and TPM were adopted to account for both the transport of these mercury forms into the model domain and their formation from background GEM within the domain.

3. Emissions

Short-term variability of mercury content in air can be potentially influenced by direct anthropogenic emissions, as well as natural emissions and anthropogenic re-emissions. For the purposes of this intercomparison, it was assumed that direct anthropogenic emissions would exert the greatest influences on observed atmospheric concentrations in the region. Travnikov and Ilyin (2005) showed that only a few percent of GEM over Europe can be accounted for European natural emission and re-emission on an annual basis. Hence only direct anthropogenic emissions were considered in the regional models. For the participating regional models, natural emission and re-emission outside the domain were partially taken into account by assumed boundary conditions and/or other model assumptions as noted above and described in more detail in Table 1. For the global models, all types of mercury emissions were considered. It is believed that natural emissions and re-emissions occur exclusively as GEM (Ebinghaus et al., 1999b).

For simulating both the 1995 and 1999 episodes, regional anthropogenic emission data for 1995 (Pacyna et al., 2001) were used, and these emissions data are referred to in this study as the "MOE" emissions inventory. The data consist of mercury emissions from point and area sources in European countries. The point and area sources were attributed to cells of a 50×50 km spatial grid. However, the Eulerian models have different horizontal resolution (from 36×36 to 150×150 km). Thus, the 50×50 km initial emissions field had to be interpolated to match the resolution for each of these models. Due to the lack of precise point-source location information, the Lagrangian models utilized virtual point sources located in the centre of each grid cell.

Emissions into each grid cell were divided into three vertical layers:<56 m, 56–136 m, and >136 m. Emissions heights and proportions of the different mercury forms for point sources were estimated on a source-by-source basis as described by Pacyna et al. (2001). Total 1995 European anthropogenic emissions in the inventory were 337 t yr^{-1}, and the emission distribution over the EMEP region for this inventory (all mercury forms, all emission levels) is presented in Fig. 1. Overall, for the entire inventory, 61% of the mercury was emitted in elemental form, 32% as RGM and 7% as TPM.

Preliminary calculations using the 1995 emission data showed that the ADOM model significantly underestimated peak concentrations of GEM measured at German sites. At the same time, the simulated and observed mean concentrations were reasonably consistent. These findings suggested that local German

Table 2
Input chemical parameters and assumptions used in the models

Data	Assumption(s) and/or estimation method
Sulphur dioxide	Daily mean values generated via air quality simulation by EMEP/MSC-W (the period from 19.06.1996 to 07.07.1996 for the first episode and the period from 23.10.1999 to 14.11.1999 for the second episode); (Iversen et al., 1990); results for 1995 were not available, so the models which utilized these data used 1999 concentrations for the 1995 episode; 50×50 km horizontal resolution; average for the boundary layer. CMAQ, GRAHM and DEHM included sulphur dioxide concentrations in simulation. ADOM used the fixed sulphur dioxide concentration value of 1 ppb.
Ozone	6-h mean ozone values generated via air quality simulation by EMEP/MSC-W from 19.06.1996 to 07.07.1996 (to be used for the first episode) and from 23.10.1996 to 14.11.1996 (to be used for the second episode); (Simpson et al., 1993); data for 1995 and for 1999 are not available; 50×50 km horizontal resolution; 5 layers within lower 4 km. CMAQ, GRAHM and DEHM included ozone concentrations in simulation. ADOM used the fixed ozone concentration value of 35 ppb.
Soot	Modelled values of soot concentrations courtesy of Dr. Trond Iversen (Iversen et al., 1998); the data are given with 150×150 km spatial resolution at 7 vertical levels up to 4 km height; the temporal resolution is 6 h. It is assumed that soot concentration outside Europe was equal to the minimum value within Europe (at corresponding vertical levels).
pH of cloud water	4.5
Chloride concentration in cloud water	2.5 mg l^{-1}
OH radical in cloud water	Midday (maximum): 10^{-12} M; night: 0
HO$_2$ radical in cloud water	Midday (maximum) concentration: 5×10^{-9} M; night-time: 0.
Chlorine concentrations over the sea	CMAQ assumes yield of 100 ppt Cl$_2$ per day within 100 m of the marine atmosphere and 5 times lower for the Baltic Sea. HYSPLIT uses Cl$_2$=5 ppt; HCl=0.005 ppt. The other models do not consider chlorine.
Mercury depletion event (MDE) in the Arctic	DEHM uses oxidation rate of Hg^0=0.25 h^{-1} during polar sunrise in the Arctic. The other models do not consider MDE.

emission sources (which are mainly responsible for elevated peak concentrations) could be underestimated by the inventory used. To check this hypothesis it was decided that for two models (ADOM and MSCE-HM) an alternative (1990) inventory (UBA: Berdowski et al., 1997) would also be used, in which emission values were much higher (463 t yr^{-1} of direct anthropogenic emissions in Europe). The ratio of mercury forms in the anthropogenic emissions was specified at the country level (Axenfeld et al., 1991). The emission sources were divided into two height categories — low (0–100 m) and high (>100 m). A significant feature of this alternative inventory is the relatively high intensity of emissions that were estimated to occur at the time in the eastern part of Germany.

The two participating large-scale models (DEHM, GRAHM) used the common European 1995 inventory described above (Pacyna et al., 2001) and supplemental emissions inventories for the remainder of their domains. The Danish DEHM model used emissions as described by Christensen et al. (2004). The Canadian GRAHM model included estimated global natural emissions and re-emissions and utilized the 1995 global 1°×1° anthropogenic emissions inventory (Pacyna and Pacyna, 2002).

4. Measurements

Measurements of concentrations of different mercury species were carried out at two German, two Swedish and one Irish monitoring stations. Locations of the stations and characteristics of surrounding areas are shown in Fig. 1. The Neuglobsow station is situated in the north-eastern part of Germany and can be influenced by industrial emission sources in Central Europe. The Zingst station is on the shore of the Baltic Sea where no major local sources are situated. However, the station can also be influenced by Central European emission sources. Swedish stations are more remote and can reflect air pollution levels over relatively clean Scandinavia. The station at Mace Head,
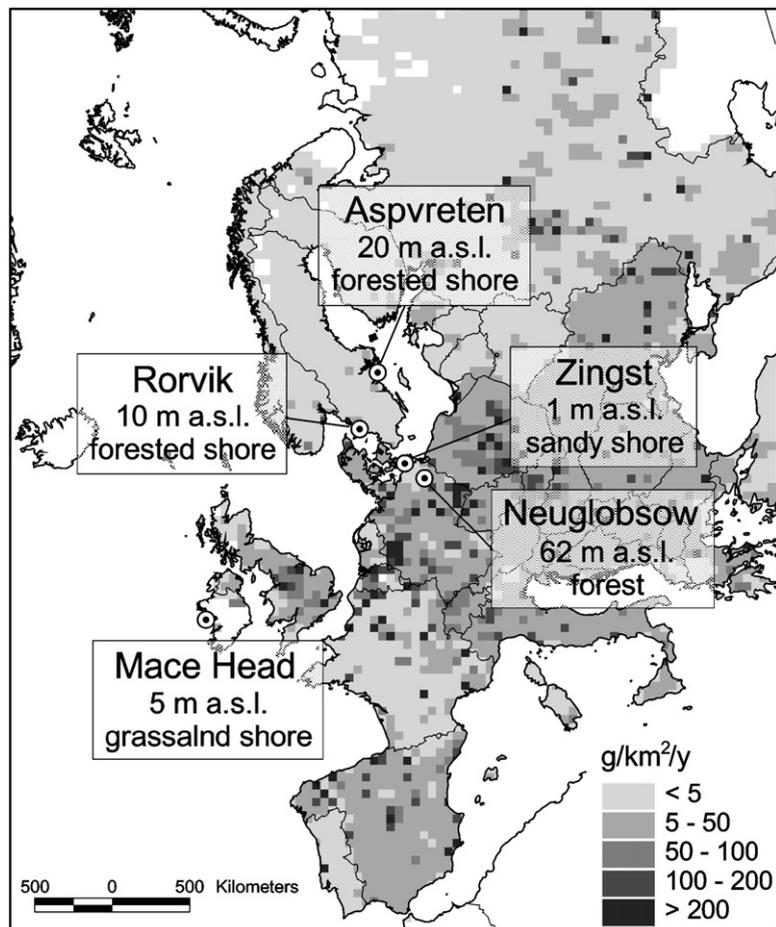

Fig. 1. Mercury anthropogenic emission distribution and locations of monitoring stations.

located in the most western part of Europe, can characterize air composition over the North Atlantic.

In the first measurement campaign (during summer 1995), only GEM concentrations were measured at the four German and Swedish stations. During the November 1999 second episode, RGM and TPM were measured all five stations, while GEM was only measured at Neuglobsow, Zingst, and Mace Head. A detailed description of all methods applied for the sampling and analysis of different atmospheric mercury species during the field measurement campaigns as well as the results obtained can be found in Schmolke et al. (1999), Munthe et al. (2001), Ebinghaus et al. (2002) and Munthe et al. (2003). The uncertainty in GEM
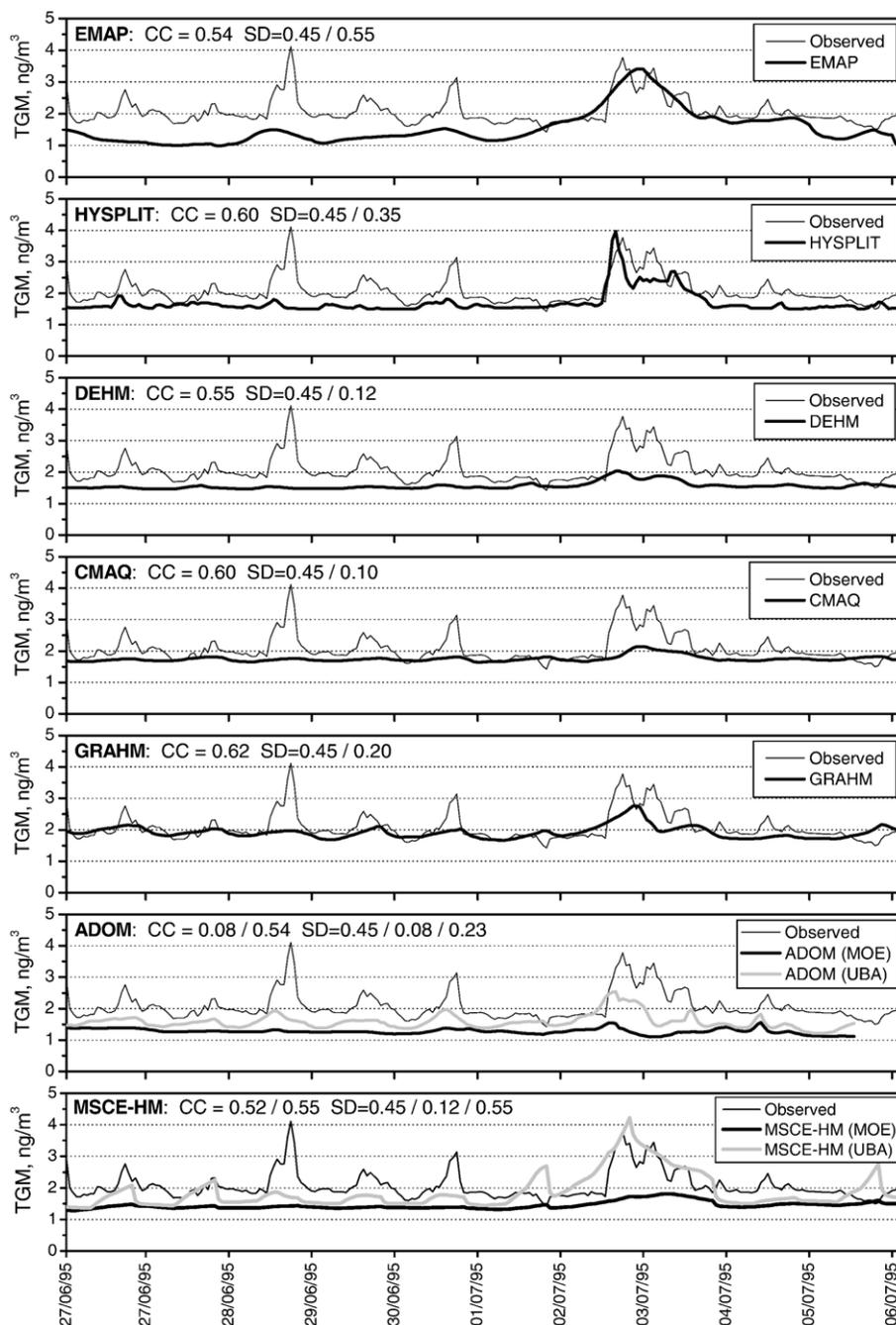

Fig. 2. GEM concentrations at Neuglobsow station: measurements and model results. CC — correlation coefficient; SD — standard deviation.

Table 3
Statistical summary of mean GEM results for two considered episodes, ng m^{-3}

Station	Parameter*	Obs**	CMAQ	ADOM	ADOM(2)***	MSCE-HM	MSCE-HM(2)***	GRAHM	DEHM	HYSPLIT	EMAP
The first episode from 26.06.1995 to 06.07.1995											
Neuglobsow	AM	**2.10**	1.75	1.28	1.59	1.44	1.88	**1.91**	1.56	1.70	1.55
	SD	**0.45**	0.10	0.08	0.23	0.12	0.55	0.20	0.12	0.35	0.55
	CC		0.60	0.08	0.54	0.52	0.55	0.62	0.55	0.60	0.54
Zingst	AM	**1.82**	1.72	1.28	1.54	1.43	1.71	**1.91**	1.52	1.58	1.31
	SD	**0.27**	0.03	0.08	0.24	0.05	0.26	0.19	0.06	0.09	0.37
	CC		0.05	0.04	0.43	−0.06	0.05	0.08	0.04	−0.03	−0.12
Rörvik	AM	**1.54**	1.70	1.29	1.35	1.42		1.77	1.49	**1.54**	1.11
	SD	**0.11**	0.02	0.10	0.10	0.03		0.15	0.03	0.04	0.11
	CC		0.09	−0.01	0.06	0.10		0.34	−0.05	0.25	0.07
Aspvreten	AM	**1.51**	1.69	1.26	1.28	1.44		1.73	**1.49**	1.53	1.19
	SD	**0.11**	0.03	0.09	0.08	0.03		0.14	0.03	0.03	0.08
	CC		−0.05	0.17	0.19	−0.48		0.28	0.18	0.23	0.13
The second episode from 01.11.1999 to 15.11.1999											
Neuglobsow	AM	**2.03**	2.32	1.20	*1.89*	2.15		**2.09**	1.80	1.91	1.35
	SD	**0.46**	0.28	0.43	*0.77*	0.34		0.29	0.19	0.38	0.56
	CC		0.50	0.21	*0.48*	0.20		0.34	0.56	0.42	0.30
Zingst	AM	**1.59**	2.13	1.07	*1.45*	2.11		1.93	**1.66**	1.85	1.17
	SD	**0.35**	0.27	0.22	*0.38*	0.22		0.25	0.16	0.30	0.36
	CC		0.65	0.24	*0.69*	0.22		0.54	0.68	0.44	0.39
Mace Head	AM	**1.88**				**1.73**		1.71	1.47	1.60	
	SD	**0.08**				0.09		0.12	0.03	0.13	
	CC					−0.66		−0.70	−0.50	−0.66	

*AM — Arithmetic Mean; SD — Standard Deviation; CC — Correlation Coefficient between measured and modelled values; **Obs — the observed values; ***using alternative (1990) emissions inventory with higher estimates.

measurement data is less than 10% (Ebinghaus et al., 1999a). Uncertainty in RGM and TPM measurement data is believed to be much higher — as much as a factor of 2 or even greater.

5. Comparison of model results against observations

5.1. Gaseous Elemental Mercury (GEM)

A comparison of modelled and measured concentrations of GEM was performed for all stations. Detailed results of this comparison are available in a technical report (Ryaboshapko et al., 2003). Here, we present summary results for all sites, but due to space limitations, we primarily discuss detailed results for just one sampling site — the Neuglobsow station. The GEM measurements at Neuglobsow for the 1995 episode showed several peaks (as high as 3–4 ng m^{-3}) which were significantly higher than the ~2 ng m^{-3} background (Fig. 2). Back-trajectory analysis can explain only one of these peaks (around July 3). During this period air masses arrived from industrialized areas of Germany, Poland, and the Czech Republic with significant emissions. Most of the models were able to capture in some way this period of peak concentrations at

Neuglobsow. The two models that utilized the UBA inventory showed better agreement using that inventory as input than with the MOE inventory. However, even using the MOE inventory, several models were also able to capture the occurrence of these peaks reasonably well.

In contrast, it can be seen that all of the models either missed completely or significantly underestimated the first large peak (~5 AM, June 29) at Neuglobsow, and two later, smaller peaks, even when the higher-emissions (UBA inventory) were utilized. Back-trajectory analysis showed that the air masses associated with these measured peaks came from the Northwest (a region that did not have major emissions in any of the inventories used). It is possible that these peaks were caused by a local source of mercury emissions which was missed in the emission inventory. Another possible cause of this discrepancy may be connected with meteorological factors. We note that practically all of those peaks occurred during nighttime or early morning. Re-emissions or natural emissions (while likely lower at night due to lower temperatures) might lead to short-term elevated concentrations in the surface layer of the atmosphere, which is generally very stable at night. Since the measurements were just at ground level, the observed peaks might have been caused by this phenomenon. After sunrise, the atmosphere generally

Table 4
Statistical summary of RGM results, pg m^{-3}

Station	Parameter*	Observation	CMAQ	ADOM	MSCE-HM	GRAHM	DEHM	HYSPLIT	EMAP
Neuglobsow	AM	**9.0**	32.1	23.6	19.6	36.3	3.5	12.9	**8.9**
	CC		0.91	0.37	0.39	0.85	0.36	0.62	0.92
Rörvik	AM	**7.4**	10.7	1.9	5.7	4.6	0.9	7.8	**7.6**
	CC		−0.23	−0.28	−0.48	0.13	−0.32	0.30	0.35
Aspvreten	AM	**5.4**	6.2	1.4	3.6	2.6	0.4	7.1	**6.0**
	CC		−0.07	−0.28	−0.43	−0.60	−0.39	−0.04	0.91
Mace Head	AM	**17.0**	–	–	2.5	0.2	0.9	13.9	–
	CC		–	–	−0.87	−0.57	0.11	−0.58	–

*AM — Arithmetic Mean; CC — Correlation Coefficient between measured and modelled values.

becomes more unstable, and the elevated concentrations near the surface are dispersed by mixing throughout the boundary layer.

Some field experiments have demonstrated that mercury air concentrations can increase 2–3 times after rain events (e.g. Lindberg et al., 1999; Wallschläger et al., 2000). However, this does not appear to be a possible explanation in this case because meteorological observations at or nearby this station did not record rain events during this period.

None of the participating models included detailed, mechanistic descriptions of natural emission and re-emission, and most of the participating models completely omitted this phenomena. This may explain the discrepancy found in the early peaks at Neuglobsow. This situation gives more impetus to the challenge of including natural and re-emissions in the models, and also suggests that additional measurements at higher elevations above ground level would be useful to evaluate and improve atmospheric models. Attempts have been made to include parameterisation of natural emission processes into regional model formulations; however, such approaches require usage of specific information like mercury concentration in soil water (Xu et al., 1999; Lin and Tao, 2003).

The GEM concentrations measured at Zingst during the 1995 episode also included several short-term peaks, primarily at night. The simulations for the Rörvik and Aspvreten stations were consistent with the GEM measurements in showing relatively small deviations from a background concentration of ∼1.5 ng m^{-3}.

Interestingly, analysis of the GEM measurements obtained at the Swedish stations reveals some cases of very low values (down to 1.1 ng m^{-3}). Such values are lower than the boundary concentrations or initial concentrations utilized in any of the models. Back trajectory analysis showed that these low concentrations were associated with air masses transported from the Arctic. It is difficult to explain these low GEM concentrations by the possible influence of Arctic mercury depletion events, because this phenomenon typically ceases by the middle of June. Moreover, comparably low GEM concentrations (1.17 ng m^{-3} at Zingst) were observed during the second episode in November, a time of the year totally untypical for such enhanced mercury depletion phenomena. If an explanation for these low concentrations can be found, more accurate boundary concentrations for mercury modelling in the EMEP and other domains might be estimated.

A statistical summary of the GEM results is shown in Table 3 where the values closest to the measurement data

Table 5
Statistical summary of TPM results, pg m^{-3}

Station	Parameter*	Observation	CMAQ	ADOM	MSCE-HM	GRAHM	DEHM	HYSPLIT	EMAP
Neuglobsow	AM	**40.2**	73.2	31.2	31.6	61.5	**45.3**	46.2	32.2
	CC		0.94	0.84	0.72	0.67	0.64	0.92	0.78
Zingst	AM	**32.8**	61.0	18.8	34.3	42.5	**31.5**	38.7	26.0
	CC		0.81	0.66	0.54	0.79	0.82	0.62	0.70
Rörvik	AM	**14.6**	45.4	15.3	14.1	40.3	19.1	24.8	**14.5**
	CC		0.88	0.87	−0.10	0.86	0.83	0.79	0.78
Aspvreten	AM	**9.5**	45.6	11.6	10.6	52.0	16.1	18.1	**10.4**
	CC		0.69	0.72	−0.22	0.59	0.70	0.76	0.53
Mace Head	AM	**13.6**	–	–	13.1	**14.1**	18.1	15.8	–
	CC		–	–	0.83	0.74	0.79	0.44	–

*AM — Arithmetic Mean; CC — Correlation Coefficient between measured and modelled values.

Table 6
Statistical summary for uncertainty assessment

Mercury species	Statistical parameter*	CMAQ	ADOM	MSCE	GRAHM	DEHM	HYSPLIT	EMAP
GEM	F_{max}	1.6	2.7	1.8	1.6	1.6	1.5	2.2
	CF_2 (%)	100	93	100	100	100	100	99
	F_{mean}	1.2	1.4	1.2	1.2	1.2	1.1	1.4
RGM	F_{max}	7	79	4900	245	41	5	19
	CF_2 (%)	48	29	19	30	15	59	90
	F_{mean}	2.2	4.9	5.1	5.7	9.1	1.9	1.6
TPM	F_{max}	38	15	33	18	13	20	16
	CF_2 (%)	29	73	59	56	78	65	62
	F_{mean}	3.1	1.8	2.3	2.2	1.7	2.0	1.9

*See the text.

are highlighted in bold. First, it is encouraging to note that for the first episode (1995) all modelling results are within ±40% of the mean observed values. Second, it can be seen that all models tend to underestimate the GEM concentrations at German stations and to overestimate them at Swedish stations. Standard deviation (SD) values show that in practically all cases the observational values vary more widely than the modelled ones. It means that the models tend to smooth away short-term variations. The most pronounced variability of the concentrations (the highest standard deviation) is characteristic of Neuglobsow — the most polluted station. For this station the models demonstrate rather high correlation with the observations. It should be noted that usage of higher anthropogenic emissions (UBA scenario) leads to better agreements with observations. The observations at Mace Head are characterized by very low variations (the lowest SD value). In this case, the models cannot demonstrate any positive correlation with the observations. Negative values reflect the fact that during the episode the observed concentrations decreased very slowly, while the modelled values rose slowly.

For the second episode (1999), one can see that the models do not tend to underestimate observations. A possible explanation for this is that the 1995 emissions inventory used may have overestimated the emissions during this 1999 episode, as anthropogenic mercury emissions were reduced between 1995 and 1999. The models using MM5-generated meteorological data (CMAQ, DEHM, and HYSPLIT) demonstrate better correlation coefficients. However, the correlation coefficients are generally low relative to the first episode.

5.2. Reactive Gaseous Mercury (RGM)

Modelling RGM concentrations represents a more challenging task, as the understanding of the physico-chemical properties and atmospheric chemistry of this mercury form is very limited. Moreover, we know very

little about the actual chemical species comprising RGM in emissions or in the atmosphere. Parameters for the estimation of wet and dry deposition for RGM are also very uncertain. To compound the problem, there is a significant measurement uncertainty for RGM (Ebinghaus et al., 1999a). Hence, it was not expected that the models would be able to closely match the observed concentrations.

Table 4 presents measured and modelled concentrations of RGM as well as correlation coefficients for each monitoring station. It can be seen that there is a broad variation in measured and modelled concentrations among the four stations. For the most polluted station (Neuglobsow), most models (all except for DEHM and EMAP) tended to overestimate the concentrations values. For the two Swedish stations (Rörvik and Aspvreten), both overestimation and underestimation took place. The highest discrepancies were found at Mace Head, where the ratio reached 2 orders of magnitude. Relatively high RGM concentrations were measured, but most of the models (except HYSPLIT) predicted very low RGM concentrations at this remote Atlantic-coast site. Potential reasons for HYSPLIT's apparent relative success include (a) reactive chlorine oxidation processes were included, in contrast to other models (except CMAQ); (b) dissolved, oxidized
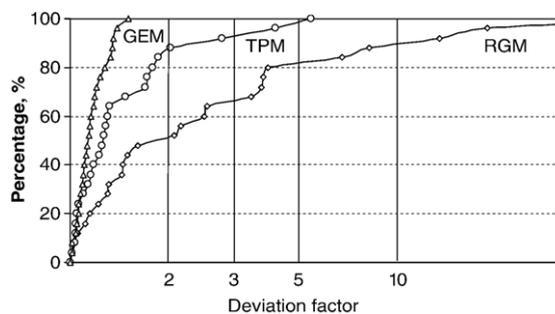

Fig. 3. Uncertainty factors vs. percentage of the data.

mercury was conditionally considered to be RGM; and (c) an assumed background RGM concentration of 5 pg m^{-3} was utilized everywhere in the model domain (in contrast to the boundary conditions assumed in the Eulerian models).

In spite of the high discrepancy between modelled and observed mean concentration values at Neuglobsow and recognition that there were relatively few data for comparison (6–7 samples), there was an encouraging correlation between observed and modelled values. At the same time three of four models demonstrate significant anti-correlation for Mace Head. Correlation coefficients for Swedish stations are mostly negative. These findings are consistent with the general understanding that that there are two independent types of RGM sources in the atmosphere — direct anthropogenic RGM emissions and *in-situ* atmospheric formation (e.g., Munthe et al., 2003). At the Central European site (Neuglobsow) anthropogenic sources prevail, and the models were able to simulate their atmospheric fate and transport reasonably well. However, at other locations, *in-situ* RGM formation was likely more important, and the simulations were less successful. This phenomenon may be most important near the sea surface and in the marine boundary layer due to elevated levels of halogen oxidants. This may explain why the RGM concentrations in marine air masses at Mace Head were generally higher than in continental air masses. The models were not able to capture this occurrence, resulting in an anti-correlation of observed and modelled results. The Swedish stations are located in coastal areas, but are also influenced by direct anthropogenic RGM emissions. Thus, both types of sources may be of similar importance at those sites.

5.3. Total Particulate Mercury (TPM)

The simulation of atmospheric particulate mercury is challenging, but its behaviour is believed to be somewhat less complex than that of RGM, e.g., insofar as chemical and phase transformations are concerned. Summaries of the TPM measurements and simulation results are given in Table 5. In general, the models were able to estimate TPM concentrations more accurately than RGM concentrations. In the case of German stations (Neuglobsow and Zingst) the mean modelling results are reasonably close to the measured values. At the Swedish stations only GRAHM and CMAQ significantly overestimate the concentrations (up to a factor of 5). Surprisingly, the four models that simulated the TPM concentrations at the "background" Mace Head station found very good agreement between model

estimates and observations. The measurements at Mace Head station showed that TPM concentrations varied from ~0 to 30 pg m^{-3}. A back-trajectory analysis confirmed that the observed variability is most likely connected to changes of air mass transport direction. Higher TPM concentrations were observed and modelled when winds were blowing from the continent.

There were a larger number of TPM measurements (9–14) than that for RGM. Hence, the estimated correlation statistics in this case may be more reliable. All the models demonstrated very high correlations with the individual observations for German stations. Reasonable agreements with the observations and significant correlation suggest that the models are able to adequately predict regional and long-range transport of particulate mercury at least at the Central European scale. High correlations obtained for Mace Head suggest that TPM concentrations in the atmosphere are affected by anthropogenic emission sources in Europe.

6. Uncertainty of modelling results

Users of modelling information need to know the degree of reliability or level of uncertainty of modelling results. As a way to characterize the ability of the different models to reproduce the observations, the *Uncertainty Factor* (F), obtained by dividing the greater value by the smaller one in each "observation–calculation" pair, is commonly used. If a series of "observed/modelled" pairs are treated, the maximum value of the factor (F_{max}) serves as a measure of the maximum inaccuracy of the modelling results, i.e. all the factors of a given series do not exceed F_{max}. The EMEP/WMO/UNEP workshop on heavy metals atmospheric transport modelling recommended that models should provide the results within a "factor of 2", i.e. annual results of regional-scale models should not be more than 2 times smaller or greater than corresponding observational results (Ilyin et al., 2000). In addition to F_{max}, there are two useful statistical characteristics, which can represent uncertainty of modelling results. First, the *Factor of Two Coverage* (CF_2) is determined as the fraction of results, which are within a factor of two. Second, the *Mean Deviation Factor* (F_{GM}) is a measure of scattering of the comparison results, and it is obtained as a geometric mean value of all "the greater to the smaller" F ratios in each compared pair.

Results of the above uncertainty assessments for each mercury form are presented in Table 6. To provide comparability with the other mercury forms, GEM daily mean values were used for calculation of the statistical parameters. The coefficients in the table represent the

totality of the results obtained for all measurement stations. The table demonstrates that all the models can predict daily mean concentrations of GEM within or very close to the recommended measure of accuracy. Essentially all individual results are within a factor of 2. The scattering of the results is also low — only ADOM and EMAP exhibit an F_{mean} value higher than 1.2. This analysis suggests that the models are generally able to simulate daily mean GEM concentrations to within a factor of two.

For RGM, it is clear that the models are generally unable to reproduce individual measurements on a sample-by-sample basis. Deviation of modelling results from the measurements can reach 2 orders of magnitude. However, while the scattering of the results is very high, the bulk of the modelling results are within a factor of 10 of the measurements. Only in the case of HYSPLIT and EMAP are more than half of the values within a factor of two. These results confirm that our knowledge of RGM's atmospheric behaviour is not sufficient, and there is a high level of uncertainty associated with the model estimation of RGM concentrations at any particular time and place.

For the TPM concentrations the models can generally reproduce the observed values for individual samples. Most results of the models (excluding CMAQ) lie within a factor of 2. The CMAQ model provides a high correlation with the observations but tends to overestimate the TPM concentrations. Overall, considering the measurement uncertainty, the performance of the models in estimating TPM is quite satisfactory. The statistical parameters discussed above can only qualitatively characterize the probability distribution of deviations of modelled results from observations. The probability curves for GEM, TPM and RGM concentrations for the totality of the participating models are presented in Fig. 3. These curves were calculated for mean concentration values during two-week episodes. Because the bulk of the uncertainty factors are lower than 5, but the highest values can reach two orders of magnitude, the ordinate of the figure is presented on a logarithmic scale.

One can see that the curves (especially for RGM and TPM) are highly non-linear. The significance of this is that a small fraction of the data exerts a strong influence on the overall uncertainty. However, the probability of obtaining such large deviations is relatively low. The figure demonstrates that in this evaluation, the models were able to predict two-week mean concentrations of GEM with uncertainty on the level of a factor of 1.35. For TPM concentrations, the probability of obtaining a result within a factor of 2.5 is more than 90%. Finally,

about 90% of the results for RGM lie within a factor of 10 and 50% lie within a factor of 2. Expressed another way, and assuming that these results are typical, any user of the mercury modelling information should understand that (a) the probability of a result for short-term (2-week) concentrations worse than factor of 2.5 for TPM and factor of 10 for RGM is about 10% and (b) for GEM, the current models can generally predict such short-term concentrations within ±35%.

7. Discussion and conclusions

One should keep in mind that this modelling experiment had several limitations. First, only anthropogenic emissions were used for the regional models, and as was discussed above, this may have led to inconsistencies between modelled and measured mercury concentrations in some cases. For example, some of the observed peaks at the two Central European stations may have been due to natural emissions or/and re-emissions from contaminated soils. The fact that any potential re-emissions were not included in the models may have contributed to the tendency of the models to under-predict the peak concentrations measured at these sites. However, two models, which considered to a certain extent contribution of natural emission and re-emission, do not differ noticeably from the others. This suggests that in this study, direct anthropogenic emissions had the largest influence on atmospheric concentrations of mercury.

It is interesting to note that the reduction of mercury emissions in Central Europe between 1990 and 1995 may be overestimated since test runs with two of the models indicate that the simulated peak concentrations show a better agreement with observations when a 1990 emission inventory is used. Given the extraordinary challenges in creating an accurate emissions inventory, the degree of reliability of the anthropogenic emission data used by all models is believed to be relatively low, both in terms of the *amount* and the *proportions* of different forms of emitted mercury. Moreover, information on any temporal variations in emissions, which can be important for short-term episodes, was not available for any of the sources in the inventory. The fact that models were able to reproduce some GEM peaks but not others may have been at least partially caused by such limitations in the emission inventory data. Uncertainty of the emission values can exceed 50% for individual grid cells. Finally, we note that the emissions inventory was specified on a 50×50 km grid. Thus, the impact of any sources within ~100–200 km of the measurement sites is unlikely to have been simulated properly.

The models themselves had many limitations. Uncertainties in the atmospheric chemistry of mercury are acknowledged to be significant. GEM concentrations are generally 1–2 orders of magnitude greater than RGM concentrations. Thus, small errors – relative to the TGM concentration – in estimating transformation rates between elemental and divalent mercury could lead to large errors in estimates of RGM concentrations. Thus, it is perhaps not surprising that the largest discrepancies between modelled and measured concentrations were found for RGM. Particulate mercury is not believed to be a particularly active agent in atmospheric chemical transformation phenomena. This may in part explain the relative success of the models in simulating TPM.

A model's spatial resolution can be an important factor in its ability to accurately simulate short-term concentration variability, particularly in regions of intense local emissions. The participating models had resolutions on the order of 50 km or greater, and this relatively coarse resolution may have led to some of the observed discrepancies.

Concentration values of gaseous elemental mercury at the inflow boundaries seem to be of importance for the participating regional scale models. There is evidence now from long-term measurements at the Irish station that average Hg^0 concentrations are higher than the boundary concentrations used with the models. In addition, the long-term measurements also show a certain seasonal variability — not accounted for in this exercise, which may have an effect on concentrations in the model domain particularly at locations relatively far from major anthropogenic sources.

Model simulations of reactive gaseous mercury (RGM) have confirmed that our knowledge about atmospheric physico-chemical processes of this mercury form is still incomplete. For example, the models tended to significantly underestimate RGM concentrations observed at the Irish west coast, perhaps suggesting that RGM formation in the marine atmosphere is more significant than that represented in the participating models.

Atmospheric models synthesize and test our current understanding. Regardless of outcome, it is therefore instructive to evaluate the performance of models in their ability to simulate real-world behaviour. The results of the study provide a quantitative example of the extent of discrepancy/agreement between measured and modelled values obtained for short-term episodes. However, it is clear that much remains to be learned about the complex atmospheric behaviour of mercury.

The comparison of the models between each other has demonstrated that they can produce similar results in spite of very different treatments of air mass transport and differences in the description of mercury atmospheric chemistry. The comparison of the modelling results with measured concentrations of three mercury forms revealed that current models of long-range transport and deposition can qualitatively – but generally not quantitatively – predict sharp variations of gaseous mercury on daily time scales. Uncertainty of GEM results can be characterized by a factor of 1.5 for daily mean concentration values and by a factor of 1.35 for two-week episodes. The models can simulate temporal variations of particulate mercury concentrations on a sample-by-sample basis with acceptable accuracy (about 90% of the results were within a factor of 2). Uncertainties in the current understanding of RGM's formation and behaviour in the atmosphere significantly limit the ability of models to quantitatively assess the atmospheric transport and deposition of this mercury form. In this study, deviations of the model results from measurements for RGM were very large, on the order of a factor of 10.

Acknowledgments

References

Axenfeld F, Munch J, Pacyna JM. Belastung von Nord- und Ostsee durch okologisch gefahrliche Stoffe am Beispiel atmospharischer Quecksilberkomponenten. Teilprojekt: Europaische Test- Emissionensdatenbasis von Quecksilber-Komponenten fur Modellrechnungen. Dornier, Report 104 02 726; 1991.

Berdowski JJM, Baas J, Bloos JPJ, Visschedijk AJH, Zandveld PYJ. The European Emission Inventory of Heavy Metals and Persistent Organic Pollutants for 1990. TNO Institute of environmental sciences, energy research and process innovation, UBA-FB report 104 02 672/03, Apeldoorn; 1997.

Bullock Jr OR, Brehme KA. Atmospheric mercury simulation using the CMAQ model: formulation description and analysis of wet deposition results. Atmos Environ 2002;36:2135–46.

Bullock OR, Brehme KA, Mapp GR. Lagrangian modeling of mercury air emission, transport and deposition: an analysis of model sensitivity to emissions uncertainty. Sci Total Environ 1998;213:1–12.

Christensen JH, Brandt J, Frohn LM, Skov H. Modelling of mercury in the Arctic with the Danish Eulerian Hemispheric Model. Atmos Chem Phys 2004;4:2251–7.

Cohen M, Artz R, Draxler R, Miller P, Poissant L, Niemi D, et al. Modeling the atmospheric transport and deposition of mercury to the Great Lakes. Environ Res 2004;95:247–65.

Dastoor AP, Larocque Y. Global circulation of atmospheric mercury: a modelling study. Atmos Environ 2004;38:147–61.

Ebinghaus R, Jennings SG, Schroeder WH, Berg T, Donaghy T, Guentzel J, et al. International field intercomparison measurements of atmospheric mercury species at Mace Head, Ireland. Atmos Environ 1999a;33:3063–73.

Ebinghaus R, Tripathi RM, Wallschlager D, Lindberg S. Natural and Anthropogenic Mercury Sources and Their Impact on the Air–Surface Exchange of Mercury on Regional and Global Scales. p. 3–50. In: Ebinghaus R, Turner RR, de Lacedra LD, Vasiliev O, Salomons W, editors. Mercury contaminated sites. Springer-Verlag; 1999b. p. 538.

Ebinghaus R, Kock HH, Coggins AM, Spain TG, Jennings SG, Temme Ch. Long term measurements of atmospheric mercury at Mace Head, Irish west coast, between 1995 and 2001. Atmos Environ 2002;36:5267–76.

Forlano L, Hedgecock IM, Pirrone N. Elemental gas phase atmospheric mercury as it interacts with the ambient aerosol and its subsequent speciation and deposition. Sci Total Environ 2000;259:211–22.

Gusev A, Ilyin I, Petersen G, van Pul A, Syrakov D. Long-range transport model intercomparison studies. Model intercomparison study for cadmium. EMEP Meteorological Synthesizing Centre EAST, report 2/2000, Moscow; 2000. 41 pp. (http://www.msceast.org.html).

Ilyin I, Munthe J, Petersen G, Ryaboshapko A. Numerical Models of Long-range Atmospheric Transport of Heavy Metals: Current State, and Direction of Further Development. WMO Report No. 136 "WMO/EMEP/UNEP Workshop on Modelling of Atmospheric Transport and Deposition of Persistent Organic Pollutants and Heavy Metals". Vol. I. Geneva, Switzerland, 16–19 November 1999, 92–111, 2000.

Ilyin I, Ryaboshapko A, Afinogenova O, Berg T, Hjellbrekke A-G, Lee DS. Lead, cadmium and mercury transboundary pollution in 2000. EMEP/MSC-E Report 5/2002, Meteorological Synthesizing Centre — East, Moscow, Russia; 2002. 131 pp. http://www.msceast.org.html.

Iversen T, Halvorsen NE, Saltbones J, Sandnes H. Calculated budgets for airborne sulphur and nitrogen in Europe. EMEP/MSC-W Report 2/90. Oslo: Norwegian Meteorological Institute; 1990.

Iversen T, Kirkevag A, Seland O. Hemispheric scale modelling of sulphate and black carbon and their direct radiative effect. In: Gryning SE, Chaumerliac N, editors. Air pollution modelling and its application, XII, vol. 22. NY: Plenum Press; 1998. p. 477–87.

Lin X, Tao Y. A numerical modelling study on regional mercury budget for eastern North America. Atmos Chem Phys 2003;3:535–48.

Lindberg SE, Zhang H, Gustin M, Vette A, Marsik F, Owens J, et al. Increases in mercury emissions from desert soils in response to rainfall and irrigation. J Geophys Res 1999;104:21879–88.

Munthe J, Wängberg I, Pirrone N, Iverfeldt A, Ferrara R, Costa P, et al. Intercomparison of methods for sampling and analysis of atmospheric mercury species. Atmos Environ 2001;35:3007–17.

Munthe J, Wängberg I, Iverfeldt Å, Lindqvist O, Strömberg D, Sommar D, et al. Distribution of atmospheric mercury species in Northern Europe: final results from the MOE Project. Atmos Environ 2003;304:S9–S20.

Pacyna EG, Pacyna JM. Global emission of mercury from anthropogenic sources in 1995. Water Air Soil Pollut 2002;137:149–65.

Pacyna EG, Pacyna JM, Pirrone N. European emissions of atmospheric mercury from anthropogenic sources in 1995. Atmos Environ 2001;35:2987–96.

Petersen G, Iverfeldt A, Munthe J. Atmospheric mercury species over Central and Northern Europe. Model calculations and comparison with observations from the Nordic Air and Precipitation Network. Atmos Environ 1995;29:47–67.

Petersen G, Bloxam R, Wong S, Munthe J, Krüger O, Schmolke SR, et al. A comprehensive Eulerian modelling framework for airborne mercury species: model development and applications in Europe. Atmos Environ 2001;35:3063–74.

Ryaboshapko A, Bullock OR, Ebinghaus R, Ilyin I, Lohman K, Munthe J, et al. Comparison of mercury chemistry models. Atmos Environ 2002;36:3881–98.

Ryaboshapko A, Artz R, Bullock OR, Christensen J, Cohen M, Dastoor A, Davignon D, Draxler R, Ebinghaus R, Ilyin I, Munthe J, Petersen G, Syrakov D. Intercomparison Study of Numerical Models for Long-Range Atmospheric Transport of Mercury. Stage II. Comparison of Modeling Results with Observations Obtained During Short-Term measuring Campaigns. June 2003. Meteorological Synthesizing Centre — East, Cooperative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe. EMEP/MSC-E Report 1/2003, Meteorological Synthesizing Centre — East, Moscow, Russia, 2003. (http://www.msceast.org.html).

Ryaboshapko A, Bullock R, Christensen J, Cohen M, Dastoor A, Ilyin I, Petersen G, Syrakov D, Travnikov O, Artz R, Davignon D, Draxler R, Munthe J, Pacyna J. Intercomparison study of atmospheric mercury models: 2. Modelling results vs. long-term observations and comparison of country atmospheric balances. Sci Total Environ 2007. doi:10.1016/j.scitotenv.2007.01.071.

Schmolke S, Schroeder WH, Munthe J, Kock HH, Schneeberger D, Ebinghaus R. Simultaneous measurements of Total Gaseous Mercury at four sites on a 800 km transect: spatial distribution and short time variability of Total Gaseous Mercury over Central Europe. Atmos Environ 1999;33:1725–33.

Shannon JD, Voldner EC. Modeling atmospheric concentrations of mercury and deposition to the Great Lakes. Atmos Environ 1995;29:1649–61.

Simpson D, Andersson-Skoeld Y, Jenkin ME. Updating the chemical scheme for the EMEP MSC-W oxidant model: current status. EMEP/MSC-W Note 2/93. Oslo: Norwegian Meteorological Institute; 1993.

Sofiev M, Maslyaev A, Gusev A. Heavy metal model intercomparison. Methodology and results for Pb in 1990. EMEP/MSC-E Report 2/1996, Meteorological Synthesizing Centre — East, Moscow, Russia; 1996. 32 pp. (http://www.msceast.org.html).

Syrakov D. On a PC-oriented Eulerian Multi-Level Model for Long-Term Calculations of the Regional Sulphur Deposition. In: Gryning SE, Schiermeier FA, editors. Air pollution modelling and its application XI, vol. 21. N.Y.: Plenum Press; 1995. p. 645–6.

Travnikov O, Ilyin I. Regional Model MSCE-HM of Heavy Metal Transboundary Air Pollution in Europe. EMEP/MSC-E Technical Report 6/2005, Meteorological Synthesizing Centre — East, Moscow, Russia; 2005. http://www.msceast.org.html.

Wallschläger D, Kock HH, Schroeder WH, Lindberg SE, Ebinghaus R, Wilken R-D. Mechanism and significance of mercury volatilization from contaminated floodplains of the Germany river Elbe. Atmos Environ 2000;34:3745–55.

Xu X, Yang X, Miller D, Helble J, Carley R. Formulation of bi-directional atmosphere–surface exchanges of elemental mercury. Atmos Environ 1999;33:4345–55.