IDŐJÁRÁS

QUARTERLY JOURNAL
OF THE HUNGARIAN METEOROLOGICAL SERVICE

CONTENTS

****

http://www.met.hu/Journal-Idojaras.php

VOL. 119 * NO. 1 * JANUARY – MARCH 2015

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service

Editorial Office: Kitaibel P.u. 1, H-1024 Budapest, Hungary
P.O. Box 38, H-1525 Budapest, Hungary
E-mail: journal.idojaras@met.hu
Fax: (36-1) 346-4669

Indexed and abstracted in Science Citation Index Expanded™ and Journal Citation Reports/Science Edition
Covered in the abstract and citation database SCOPUS®

Subscription by mail:
IDŐJÁRÁS, P.O. Box 38, H-1525 Budapest, Hungary
E-mail: journal.idojaras@met.hu

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 1, January – March, 2015, pp. 1–21

Expected changes in mean seasonal precipitation and temperature across the Iberian Peninsula for the 21st century

Joan Ramon Coll[1], **Philip D. Jones**[2, 3], and **Enric Aguilar**[1]

[1] *Centre for Climate Change, Geography Department, Rovira i Virgili University, Av. Remolins 13-15, 43500 Tortosa, Spain.*

[2] *Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK.*

[3] *Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah 21589, Saudi Arabia*

**Corresponding author E-mail: joanramon.coll@urv.cat*

(Manuscript received in final form April 24, 2014)

Abstract—Three different regional climate models (DMI-HIRHAM5, HadRM3, and KNMI-RACMO2) driven by ERA-40 reanalysis and also driven by global climate models (GCMs) obtained from the EU-Ensembles project have been compared to observed data over the Iberian Peninsula (IP) to assess the accuracy of simulated precipitation and temperature. KNMI-RACMO2 and DMI-HIRHAM5 were the best models for accurately simulating precipitation and temperature, respectively, although large uncertainties still affect their simulations. The same RCM simulations driven by GCMs have been used to project the seasonal expected changes in precipitation and temperature for the periods 2011–2050 and 2051–2090 relative to 1961–2000 under the A1B climate change scenario. From the results, a clear decrease in mean precipitation is expected in most IP for spring, summer, and autumn, but no clear signal was found in winter. Moreover, future projections showed a large increase in mean temperatures in all seasons being more evident in the interior of the IP especially in summer. The decrease in mean precipitation and the increase in mean temperature projected for the IP, could worsen current drought conditions especially for the second half of the 21st century.

Key-words: regional climate models, Iberian Peninsula, precipitation, temperature, model accuracy validation, future projections, drought

1. Introduction

Future climate will include simultaneous changes in temperature and precipitation for many regions of the world. For example, in the Iberian Peninsula (IP), the observed and projected increase in temperature is expected to be accompanied by a decrease in precipitation (*IPCC,* 2007; *Sousa et al.,* 2011; *IPCC,* 2012). This combination would have a negative impact on water availability.

Previous studies have confirmed a warming trend in the 20th century across the IP using observed data (*Brunet et al.,* 2006; *Brunet et al.,* 2007), while precipitation patterns showed a high inter-annual variability, but appreciable changes have not still been identified in annual precipitation totals (*Barrera-Escoda,* 2008; *CLIVAR*, 2010).

Being consistent with the observed trends, climate models project a large increase in temperatures, but also a future decrease in precipitation of roughly 20% in southern Europe by the end of 21st century (*IPCC,* 2007) including the whole IP (*Sánchez*, 2009; *Gómez-Navarro et al.,* 2010; *Rodríguez-Puebla* and *Nieto,* 2010; *Vicente-Serrano et al.,* 2011; *Jerez et al.,* 2012; *Jerez* and *Montavez*, 2012). Again, this warming/drying combination implies an increase of drought conditions over the wider Mediterranean region (*Blenkinsop* and *Fowler,* 2007; *Mariotti et al.,* 2008; *Dai*, 2011 and 2012; *IPCC,* 2012) and also over the IP (*Beniston et al.,* 2007; *Rodríguez-Puebla* and *Nieto,* 2010;; *Sanchez et al.,* 2012). These consistent results, obtained through the analysis of model output, have to be put in the context of the large uncertainties related to the reliability of model simulations affecting the projections of trends in temperature, precipitation, and drought conditions for the coming century (*Blenkinsop* and *Fowler,* 2007; *Sheffield* and *Wood*, 2008; *Rammukainen*, 2010; *Mishra* and *Singh*, 2011; *IPCC,* 2012; *Dai*, 2012).

This article focuses on investigating the capability of three regional climate models to correctly reproduce future temperature and precipitation in the IP. We used the outputs from the RCMs driven by ERA-40 reanalysis for the period 1958–2002 to make comparisons with observed data and the same RCMs driven by associated GCMs for future projections covering the period 1951–2100. We approach this analysis by comparing climate model output belonging to the A1B climate change scenario to an overlapping observational dataset of temperature and precipitation.

2. Data

Monthly simulated temperature and precipitation data from three regional climate models (RCMs) at 25 km resolution were obtained from the EU-Ensembles project (available at http://www.ensembles-eu.org). DMI-HIRHAM5 produced by the Danish Meteorological Institute (DMI) (*Christensen* and

Christensen, 2007), HadRM3 using the HC-Q0 (normal sensitivity) developed by the Hadley Centre for Climate Prediction and Research (HC) (*Collins et al.,* 2010), and KNMI-RACMO2 produced by the Royal Netherlands Meteorological Institute (KNMI) (*van Meijgaard et al.*, 2008) are the RCMs selected to assess changes in precipitation and temperature over the IP along the 21st century. These RCMs were selected because they are regarded as the best-performing models at handling precipitation variability in other European regions, with RACMO2 performing best overall for the UK (*Simpson*, 2011, *van der Linden* and *Mitchell*, 2009; *Christensen et al.,* 2010; *Kjellström et al.,* 2010). We used the outputs from the RCMs driven by ERA-40 reanalysis for the period 1958–2002 to make comparisons with observed data and the same RCMs driven by associated global climate models (GCMs) for future projections covering the period 1951–2100. The ERA-40 reanalysis is produced by the European Centre for Medium-Range Weather Forecasting (ECMWF), and it is based on observed data such as conventional observations or satellite data among others. Temperature from the ERA-40 reanalysis is well simulated in the Northern Hemisphere when compared with mean observed sea-level pressure and geopotential 500hPa temperatures (*ECMWF*, 2004), and precipitation is well-handled in the Northern Hemisphere continents (*Bosilovich et al.*, 2008). In this study, KNMI-RACMO2 and DMI-HIRHAM5 are coupled with ECHAM5-r3 as a GCM, while HadRM3 is associated with HadCM3. The A1B climate change scenario was chosen to project the expected changes in precipitation and temperature over the IP as it represents a medium greenhouse gases forcing to the climate system according to IPCC AR-4 (*IPCC,* 2007).

The observed data has been extracted from the Monthly Iberian temperature and precipitation series (MITPS, *Fig. 1*). The MITPS dataset updates to 2010 the Spanish daily adjusted temperature/precipitation series (SDATS/SDAPS) and adds to its 22 stations two new data points to represent the western part of the IP (Portugal). The SDATS – and its MITPS update - was quality controlled (QC) following *Aguilar et al.,* (2002) and were homogenized by the Centre for Climate Change (C3) (*Brunet et al.,* 2006; *Brunet* et al., 2007) applying the standard normal homogeneity test (SNHT) (*Alexandersson* and *Moberg*, 1997). The two Portuguese series (Lisboa and Porto) have been subjected to quality control procedure of raw data and tested using homogenization procedure based on the standard normal homogeneity test, to detect and adjust most prominent inhomogeneities on a monthly scale.

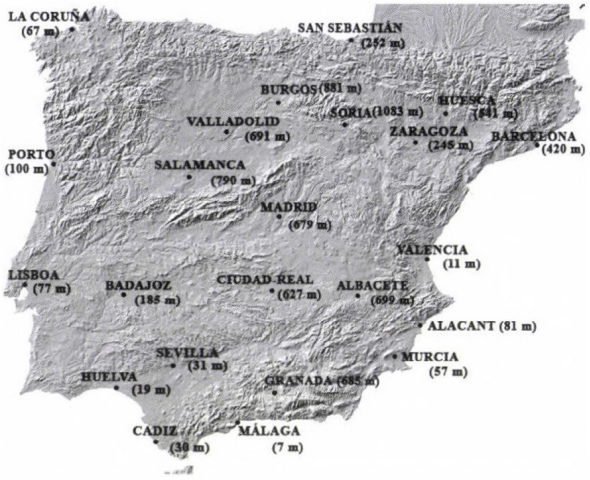

Fig. 1. MITPS station map: the closest city to each meteorological station is shown with the elevation (in brackets) of the last segment of record.

3. Methodology

The closest RCM grid box to each meteorological station has been selected to obtain the same geographical distribution of the Iberian simulated data according to observed data (MITPS) (*Fig. 1*). Monthly simulated precipitation and temperature data were obtained from DMI-HIRHAM5, HadRM3, and KNMI-RACMO2 model outputs driven by ERA-40 reanalysis and those driven by GCMs for each location. The RCMs outputs driven by ERA-40 reanalysis and by GCMs have been compared with the observed MITPS for the 1961–2010 period. All RCMs are more consistent when they are driven by ERA-40 reanalysis than by GCMs regarding precipitation and temperature simulation in the IP. Accumulated daily precipitation totals and mean temperature from observed and simulated data were totaled for Northern Hemisphere winter (DJF), spring (MAM), summer (JJA), and autumn (SON) seasons for the entire period. Based on the procedure applied by *Simpson* (2011), we computed the seasonal differences in precipitation totals, mean temperature, and the ratio between the standard deviations of simulated and observed data. Modeled and observed datasets have also been compared by computing the Pearson product-moment correlation coefficient and the root mean square error (RMSE). Finally, Kolmogorov-Smirnov test (K-S test) has been applied to evaluate the similarity between both statistical distributions.

Once the performance of the three RCMs is evaluated, simulations of seasonal precipitation and temperature from the RCMs driven by GCMs have been used to project the mean expected changes in precipitation and temperature over the IP for the periods 2011–2050 and 2051–2090 relative to 1961–2000 at seasonal time-scales under the A1B climate change scenario.

4. Results

4.1. Validation of the accuracy of precipitation simulations

The analysis of temperature and precipitation simulations showed quite different results, depending both on the model, the season of the year, and the different regions of the IP.

Precipitation simulations driven by ERA-40 tended to overestimate the observed data in central and north-western IP and also in the Ebro basin, especially in winter and spring (*Fig. 2*). Underestimated precipitation was found in the south, south-western IP and in the Mediterranean coast during all seasons, being more evident in summer and autumn. KNMI-RACMO2 showed the smallest deviations compared to observed data over the IP during all seasons except in summer, when it underestimated by between 20% and 40% of mean precipitation over most of the IP. HadRM3 produced overestimates of around 20–40% of mean precipitation in central, north-western area and in the Ebro basin, especially in spring and summer, and underestimated by 40% of mean precipitation in south-western areas and in the Mediterranean region for summer and autumn. DMI-HIRHAM5 showed large overestimates greater than 60% of mean precipitation in the Ebro basin and in the Sierra Nevada for winter and summer, respectively, and large underestimates towards 20–60% of mean precipitation for most of the IP, with higher anomalies located in the western and south-western area during all seasons.

Precipitation simulations from DMI-HIRHAM5 and KNMI-RACMO2 driven by ECHAM5-r3 and HadRM3 coupled with HadCM3 have also been compared with observed MITPS to check how the RCM simulations can be altered when they are driven by GCMs (*Fig. 3*). Large differences have been detected in mean seasonal precipitation from DMI-HIRHAM5 and KNMI-RACMO2 when they are driven by ERA-40 or by their associated GCMs (*Figs. 2* and *3*), but the simulations are closer in the case of HadRM3 in all seasons. All models have a tendency to overestimate mean precipitation in winter, spring, and autumn, while underestimates were found in summer. DMI-HIRHAM5 produced large overestimates greater than 80% of mean precipitation in central, north, and north-western IP for winter, spring, and autumn, while underestimates towards 20–40% of mean precipitation were focused in the south-western area and in the Mediterranean region for summer. HadRM3 and KNMI-RACMO2 overestimated by between 40–80% of mean precipitation in the Ebro basin and in the north-western IP mainly in winter and spring while underestimates around 40% of mean precipitation were primarily in the Mediterranean region for spring and summer. Note that large underestimates of greater than 60% of mean precipitation have been identified from KNMI-RACMO2 over most of the IP in summer.

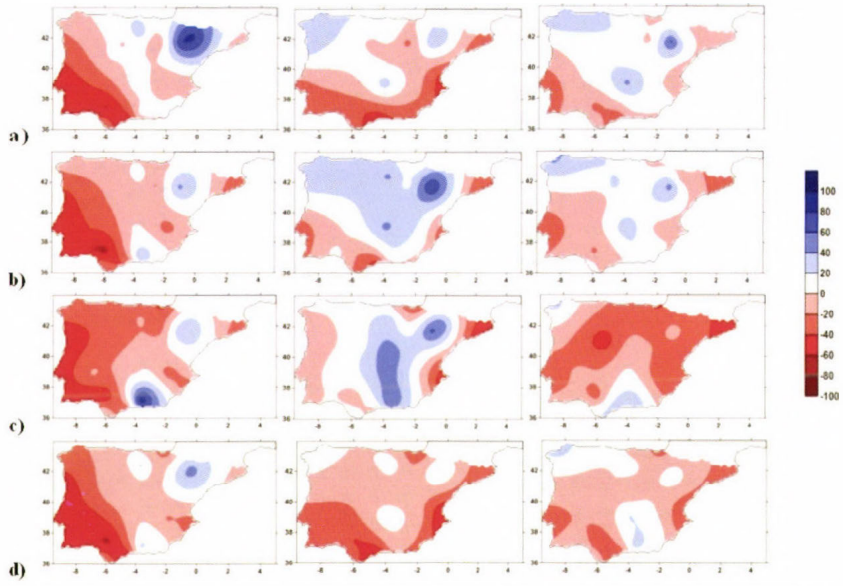

Fig. 2. Differences (in %) between simulated and observed mean seasonal precipitation totals (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 (left), HadRM3 (middle), and KNMI-RACMO2 (right), all driven by ERA-40 reanalysis.

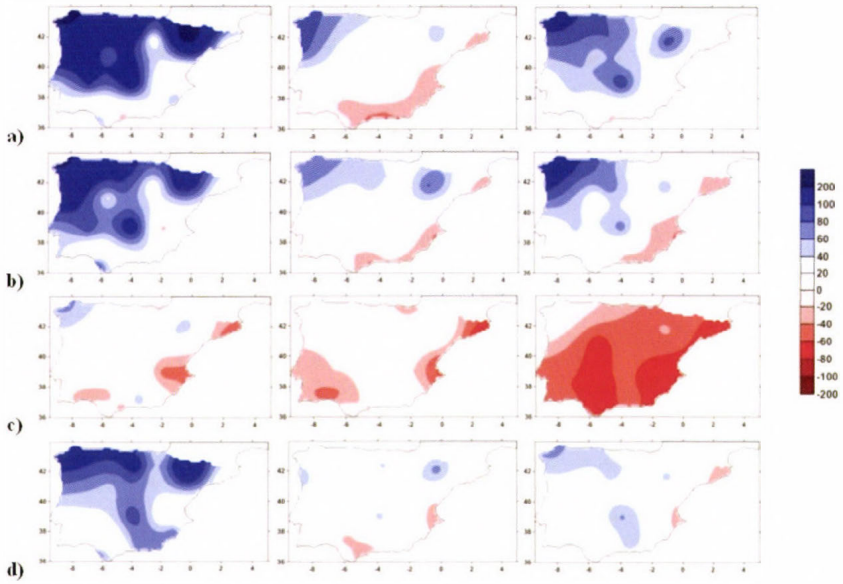

Fig. 3. Differences (in %) between simulated and observed precipitation seasonal totals (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 driven by ECHAM5-r3 (left), HadRM3 driven by HadCM3 (middle), and KNMI-RACMO2 driven by ECHAM5-r3 (right).

Simulated seasonal precipitation variability has been assessed from the ratio of standard deviations between mean simulated and observed precipitation from the three RCMs driven by ERA-40 reanalysis (*Fig. 4*). All RCMs exceeded 0.5 standard deviations for all seasons over most of the IP, while the ratio was greater than 1 standard deviation in some regions located in the north and central IP, especially for winter and spring. Otherwise, large differences in standard deviations could be appreciated from all RCMs driven by GCMs during all seasons (*Fig. 5*). All RCMs exceeded 1 standard deviation between mean simulated and observed precipitation over most of the IP, especially in spring and autumn.

The best correlations from the three RCMs driven by ERA-40 reanalysis between simulated and observed data were found in winter (*Fig. 6*), and the worst ones in summer with KNMI-RACMO2 the best-performing and DMI-HIRHAM5 second. The correlations were lower in the Mediterranean region than for the rest of the IP, especially in winter, summer, and autumn. Finally, the results from RMSE and K-S test (*Tables 1* and *2*) showed better fits between simulated and observed data using KNMI-RACMO2 during all seasons over most of the IP than DMI-HIRHAM5, which fitted better in the Mediterranean region.

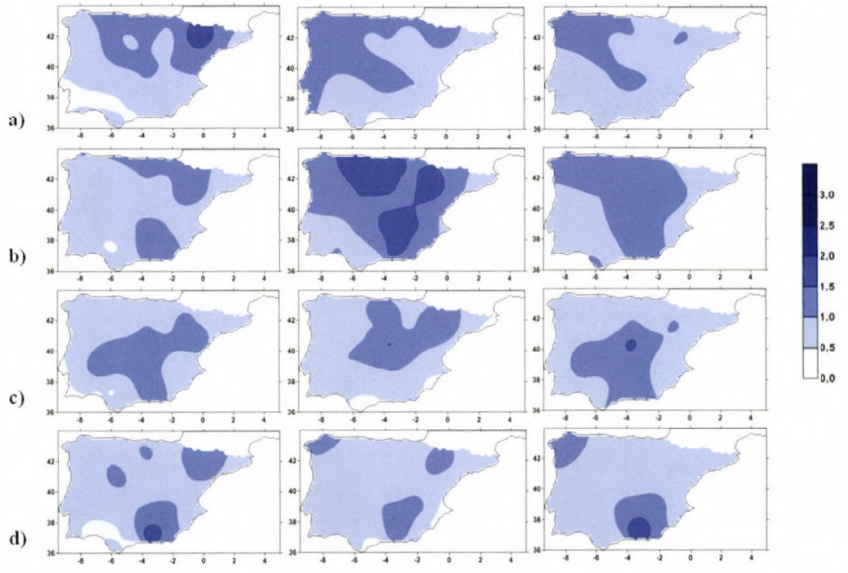

Fig. 4. Ratio of standard deviations between simulated and observed seasonal precipitation totals (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 (left), HadRM3 (middle), and KNMI-RACMO2 (right), all driven by ERA-40 reanalysis.

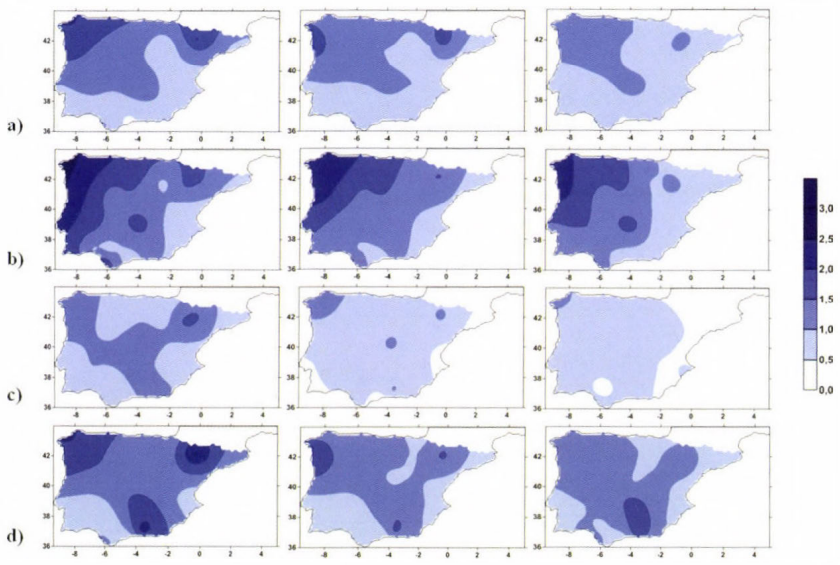

Fig. 5. Ratio of standard deviations between simulated and observed seasonal precipitation totals (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 driven by ECHAM5-r3 (left), HadRM3 driven by HadCM3 (middle), and KNMI-RACMO2 driven by ECHAM5-r3 (right).

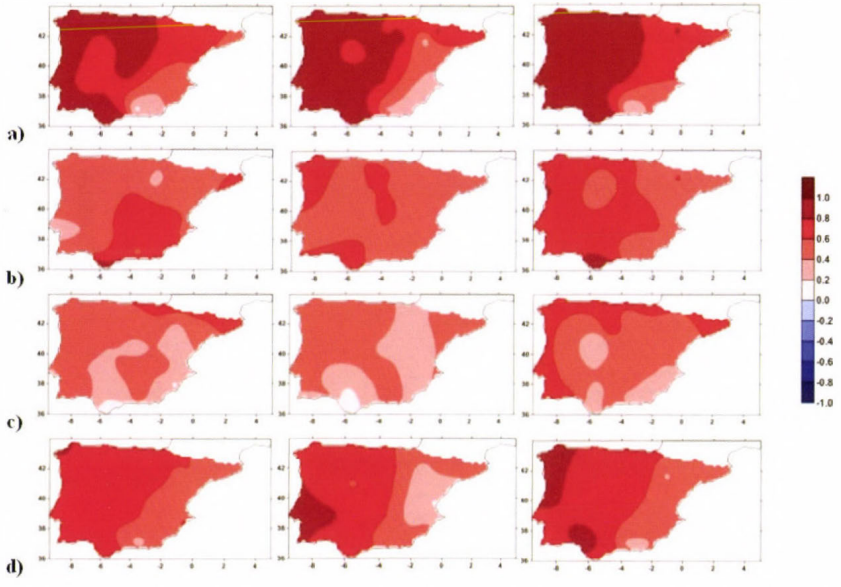

Fig. 6. Pearson Product-Moment Correlation coefficient between simulated and observed seasonal precipitation totals (MITPS) in the IP for winter (DJF); a), spring (MAM); b), summer (JJA); c), and autumn (SON); d) using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 (left), HadRM3 (middle), and KNMI-RACMO2 (right), all driven by ERA-40 reanalysis. Correlations greater than 0.41 are statistically significant at the 99% level.

Table 1. Number of the best fitted RMSE values and K-S distances between simulated and observed precipitation and temperature for each season using the common period 1961–2000 for the 23 locations spread over the IP. The model outputs are derived from DMI-HIRHAM5, HadRM3, and KNMI-RACMO2, all driven by ERA-40 reanalysis. Values in bold refer to the maximum number of the best fitted values for each goodness of fit test and for the three RCMs assessed

		DMI-HIRHAM5		HadRM3		KNMI-RACMO2	
	Seasons	**N° RMSE**	**N° K-S distances**	**N° RMSE**	**N° K-S distances**	**N° RMSE**	**N° K-S distances**
Precipitation	Winter (DJF)	4	**10**	3	3	**16**	**10**
	Spring (MAM)	10	7	0	6	**13**	**10**
	Summer (JJA)	4	5	6	7	**13**	**11**
	Autumn (SON)	5	5	3	4	**15**	**14**
Temperature	Winter (DJF)	**14**	**14**	3	3	6	6
	Spring (MAM)	**15**	**14**	4	6	4	3
	Summer (JJA)	4	4	6	5	**13**	**14**
	Autumn (SON)	**11**	8	7	**11**	5	4

Table 2. Averaged RMSE values and averaged K-S distances between simulated and observed precipitation and temperature over the whole IP for each season using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5, HadRM3, and KNMI-RACMO2, all driven by ERA-40 reanalysis. Values in bold refer to the best fitted values for each goodness of fit test and for the three RCMs assessed

		DMI-HIRHAM5		HadRM3		KNMI-RACMO2	
	Seasons	**RMSE**	**K-S distances**	**RMSE**	**K-S distances**	**RMSE**	**K-S distances**
Precipitation	Winter (DJF)	92.69	0.29	81.45	0.28	**67.25**	**0.23**
	Spring (MAM)	69.33	0.35	77.63	0.32	**59.19**	**0.26**
	Summer (JJA)	46.02	0.35	48.91	0.33	**43.83**	**0.31**
	Autumn (SON)	97.19	0.33	90.85	0.31	**78.40**	**0.25**
Temperature	Winter (DJF)	**1.37**	**0.50**	1.65	0.58	1.76	0.58
	Spring (MAM)	**1.05**	**0.40**	1.27	0.46	1.52	0.57
	Summer (JJA)	1.65	0.59	1.88	0.59	**1.34**	**0.49**
	Autumn (SON)	**1.13**	0.43	1.25	**0.42**	1.57	0.61

From the results obtained above, KNMI-RACMO2 is the most suitable RCM for simulating precipitation in the IP when driven by ERA-40 data and by GCMs, although large uncertainties in mean precipitation should be appreciated for summer. All RCMs, including KNMI-RACMO2, are still affected by uncertainties in mean precipitation that have to be taken into account for future projections especially when the RCMs are driven by GCMs.

4.2. Validation of the accuracy of temperature simulations

The outputs from the RCMs driven by ERA-40 reanalysis and driven by GCMs for temperature were also compared with observed MITPS for each location applying the methodology described earlier for the common period 1961–2000 at seasonal time-scale.

Simulations of the RCMs driven by ERA-40 showed a bias in mean seasonal temperature relative to observed MITPS towards overestimates in north-eastern IP for winter and autumn and central IP, especially in summer (*Fig. 7*). Additionally, underestimates were detected in the north and north-western IP, especially in winter, spring, and autumn and in the south-eastern IP in summer. DMI-HIRHAM5 showed the smallest anomalies of 1 °C of mean simulated temperature in most of the IP during all seasons with HadRM3 the second best. DMI-HIRHAM5 overestimated between 2–3 °C of mean temperature mainly in north-eastern IP for winter and autumn and greater than 3 °C in central IP for summer. Underestimates of between 2–3 °C of mean temperature were focused in the north and north-western area in winter, spring, and autumn, while in the south-eastern IP, similar underestimates were found for summer. HadRM3 produced overestimates of between 3–4 °C of mean temperature in the central and in north-eastern corner of the IP (around 2 °C) in winter and summer, respectively, while underestimates of between 1–3 °C were identified in the central, north, and north-western areas during winter, spring, and autumn. KNMI-RACMO2 showed overestimates of around 2 °C for mean temperature in the central IP only for summer, but underestimates were detected in the most of the IP, but were more prominent (between 3–4 °C) in the north-western area in winter, spring, and autumn.

Additionally, temperature simulations from the RCMs coupled with associated GCMs were compared with observed MITPS in order to see whether RCM simulations are affected when they are driven by GCMs (*Fig. 8*). Some differences have been detected in mean seasonal temperature from DMI-HIRHAM5 and HadRM3 when they are driven by ERA-40 or by their associated GCMs (*Figs. 7* and *8*), but the simulations fitted better in the case of KNMI-RACMO2 driven by the associated GCM than those driven by ERA-40 in all seasons. All models have a tendency to overestimate by between 1–2 °C of mean temperature in the Mediterranean region in winter, while underestimates of between 1–2 °C for mean temperature were found in most of the IP, but were

more prominent in the north-western area during winter, spring, and autumn. DMI-HIRHAM5 showed overestimates of around 2–3 °C in the Mediterranean region for winter and autumn, while the underestimates of 2–3 °C were located in the central, north, and north-western parts of the IP in spring and summer. HadRM3 and KNMI-RACMO2 produced similar underestimates of 2–3 °C over most of the IP, mainly located in the north-western corner during winter, spring, and autumn.

Seasonal temperature variability has been figured out from the ratio of standard deviations between simulated and observed average temperatures from the three RCMs driven by ERA-40 reanalysis (*Fig. 9*). All RCMs exceeded 0.5 standard deviations over the IP for all seasons being greater than 1 standard deviation over most of the IP according to HadRM3. The seasonal temperature variability between simulated and observed data from the RCMs driven by GCMs produced similar differences than the simulations driven by ERA-40 (*Fig. 10*). All RCMs exceeded 0.5 standard deviations over the whole IP during all seasons with a ratio greater than 1 standard deviation for spring, summer, and autumn according to HadRM3.

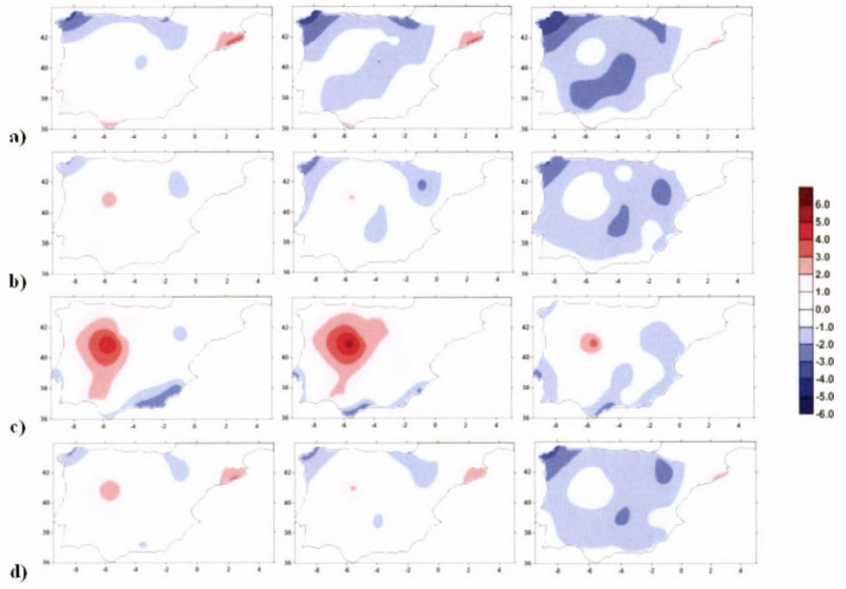

Fig. 7. Differences (in °C) between simulated and observed seasonal average temperatures (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 (left), HadRM3 (middle), and KNMI-RACMO2 (right), all driven by ERA-40 reanalysis.

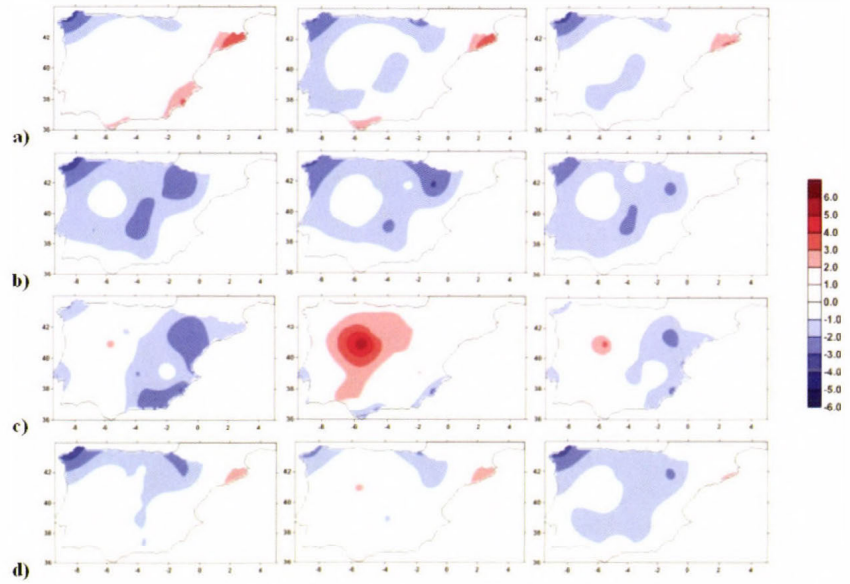

Fig. 8. Differences (in °C) between simulated and observed seasonal average temperatures (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 driven by ECHAM5-r3 (left), HadRM3 driven by HadCM3 (middle), and KNMI-RACMO2 driven by ECHAM5-r3 (right).

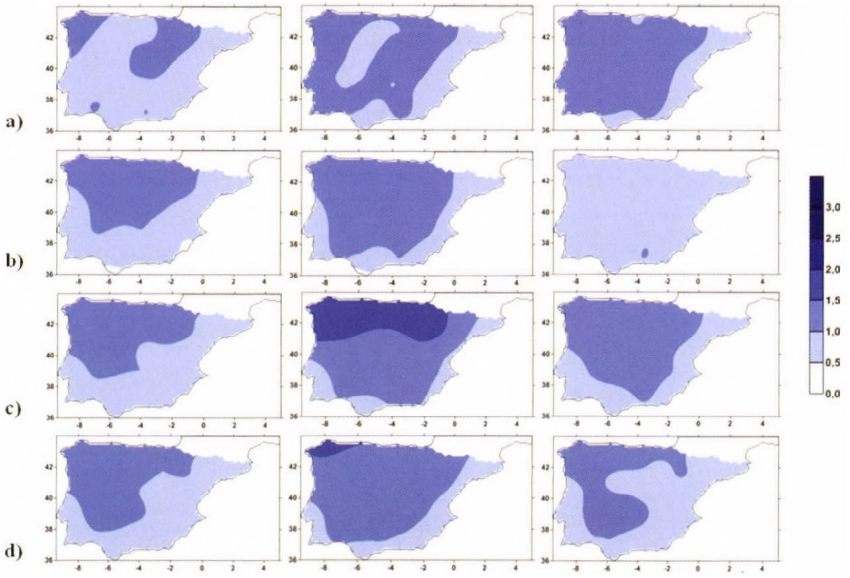

Fig. 9. Ratio of standard deviations between simulated and observed seasonal average temperatures (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 (left), HadRM3 (middle), and KNMI-RACMO2 (right), all driven by ERA-40 reanalysis.

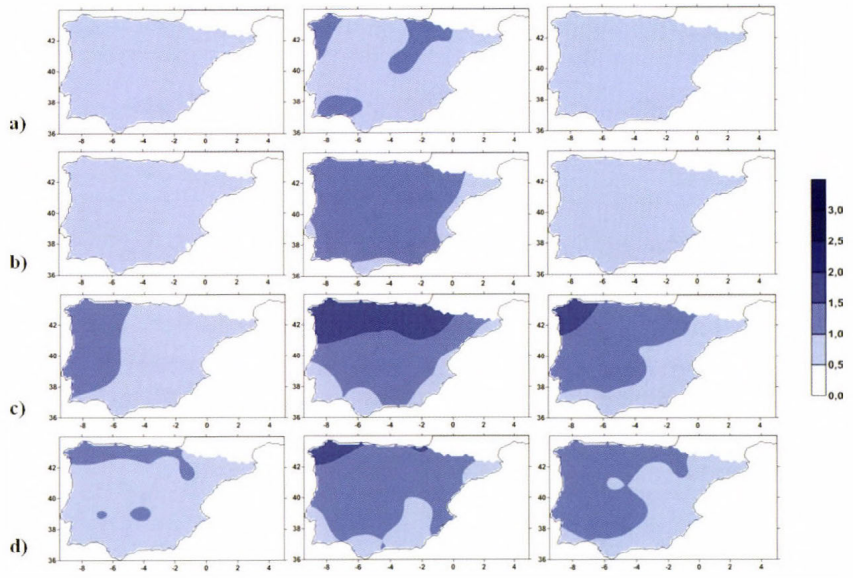

Fig. 10. Ratio of standard deviations between simulated and observed seasonal average temperatures (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 driven by ECHAM5-r3 (left), HadRM3 driven by HadCM3 (middle), and KNMI-RACMO2 driven by ECHAM5-r3 (right).

The three RCMs driven by ERA-40 reanalysis showed very high correlations between simulated and observed data during all seasons with DMI-HIRHAM5 the best-performing overall, although the correlations are lower in the Mediterranean region than in the rest of the IP, especially in summer and autumn (*Fig. 11*). Finally, the results from RMSE and K-S test showed a better fit between simulated and observed data using DMI-HIRHAM5 during winter and spring over most IP, while KNMI-RACMO2 fitted better in summer (see *Tables 1* and *2*).

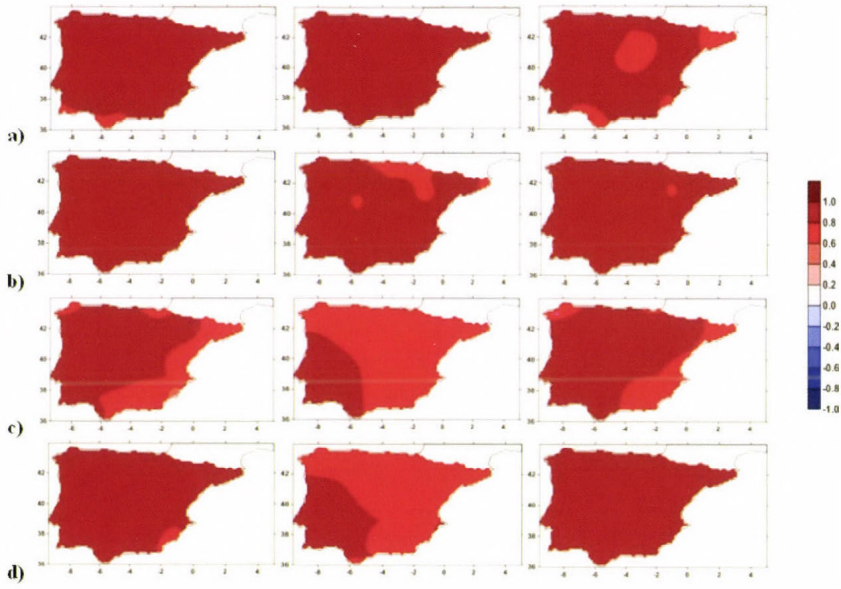

Fig. 11. Pearson product-moment correlation coefficient between simulated and observed seasonal average temperatures (MITPS) in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c)**, and autumn (SON); **d)** using the common period 1961–2000. The model outputs are derived from DMI-HIRHAM5 (left), HadRM3 (middle), and KNMI-RACMO2 (right), all driven by ERA-40 reanalysis. Correlations greater than 0.34 are statistically significant at the 99% level.

From the measures considered above, DMI-HIRHAM5 is the best RCM of the three tested for the IP for simulating temperature, although large uncertainties are affecting RCMs, especially when they are driven by GCMs. These uncertainties should be considered when projecting temperature over the IP along the 21st century.

4.3. Projected changes in mean seasonal precipitation

The outputs from DMI-HIRHAM5 and KNMI-RACMO2 driven by ECHAM5-r3 and HadRM3 driven by HadCM3 were used to assess the projected changes in mean seasonal precipitation in the IP for the periods 2011–2050 (*Fig. 12*) and 2051–2090 (*Fig. 13*) relative to 1961–2000 under the A1B climate change scenario.

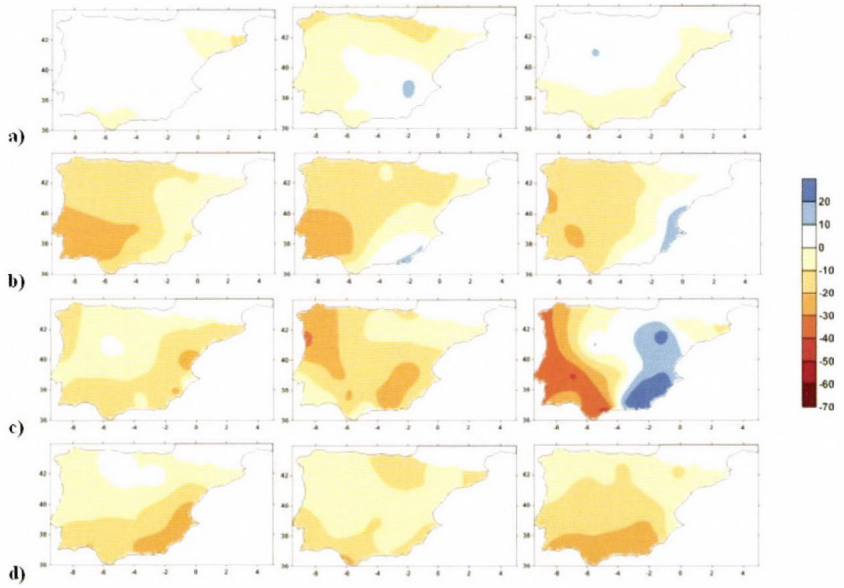

Fig. 12. Seasonal precipitation change (in %) projected for the period 2011–2050 relative to 1961–2000 in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c),** and autumn (SON); **d)** using the model outputs derived from DMI-HIRHAM5 driven by ECHAM5-r3 (left), HadRM3 driven by HadCM3 (middle), and KNMI-RACMO2 driven by ECHAM5-r3 (right).

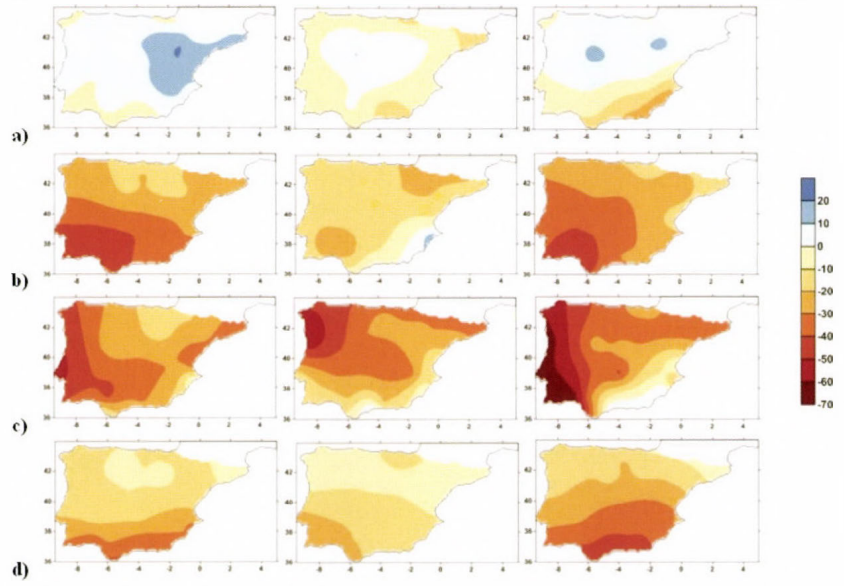

Fig. 13. Seasonal precipitation change (in %) projected for the period 2051–2090 relative to 1961–2000 in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c),** and autumn (SON); **d)** using the model outputs derived from DMI-HIRHAM5 driven by ECHAM5-r3 (left), HadRM3 driven by HadCM3 (middle), and KNMI-RACMO2 driven by ECHAM5-r3 (right).

Model outputs agree on a future decrease in precipitation over most of the IP during the 21st century for spring, summer, and autumn, although no clear signal was found for winter. All RCM simulations projected a clear decrease of 10–20% in mean precipitation for spring, summer, and autumn across most of the IP for the period 2011–2050 relative to 1961–2000 (*Fig. 12*). DMI-HIRHAM5 and KNMI-RACMO2 outputs projected similar decrease in mean precipitation of 30–50% for spring and summer over most of the IP by the 2051–2090 period, while a decline of 20–30% was detected in autumn (*Fig. 13*). HadRM3 simulations showed smaller decreases in mean precipitation of 10–20% in spring and autumn, and around 20–30% increases in summer for the mid-late century. No clear signal was found in mean winter precipitation from all the RCM outputs for both time-periods (*Figs. 12* and *13*).

In particular, DMI-HIRHAM5 and KNMI-RACMO2 outputs agree on an increase in mean winter precipitation of 5–10% in central, north, and north-western IP, and a decrease of 10% in the south, south-eastern, and in the Mediterranean region for the period 2011–2050 declining by between 10–20% by the 2051–2090 period (*Figs. 12* and *13,* respectively). All RCM outputs showed a decrease in mean spring precipitation of 10–20% for most of the IP for the period 2011–2050, although HadRM3 and KNMI-RACMO2 produced precipitation increases of 10% in some areas of the Mediterranean and south-eastern coasts of the IP. An evident spring precipitation decrease of 20–40% has been projected from all RCMs for the whole IP by the 2051–2090 period being more extreme in the south and south-western areas than in the central and in the north. DMI-HIRHAM5 produced a clear decrease of 10–20% of mean summer precipitation over most of the IP for the period 2011–2050, but declining to 20–50% by the 2051–2090 period, especially in the west and south-western area. HadRM3 showed a decline of 10–30% (2011–2050) and of 20–60% in mean summer precipitation more evident in the north and north-western IP than in the south-eastern and south-western areas. KNMI-RACMO2 projected a decrease of 10–40% in summer precipitation (2011–2050) in the west and south-western area becoming 20–70% over most IP by the 2051–2090 period. Despite this, precipitation increases were found from KNMI-RACMO2 outputs in the south-eastern area of 10–60% (2011–2050) and of 5% (2051–2090). Finally, all RCM simulations showed a clear decline in mean autumn precipitation of 10–20% (2011–2050) and of 10–40% (2051–2090) for the whole IP being more extreme in the southern than in the northern area.

4.4. Projected changes in mean seasonal temperature

The model outputs associated with their GCMs were also used to assess the projected changes in mean seasonal temperature for the IP using the periods 2011–2050 (*Fig. 14*) and 2051–2090 (*Fig. 15*) relative to 1961–2000 under the A1B climate change scenario.

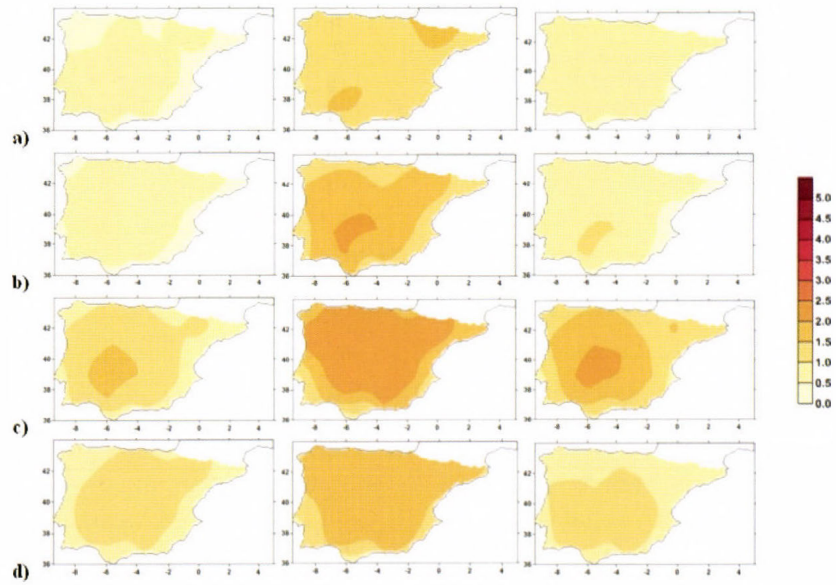

Fig. 14. Seasonal temperature change (in °C) projected for the period 2011–2050 relative to 1961-2000 in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c),** and autumn (SON); **d)** using the model outputs derived from DMI-HIRHAM5 driven by ECHAM5-r3 (left), HadRM3 driven by HadCM3 (middle), and KNMI-RACMO2 driven by ECHAM5-r3 (right).

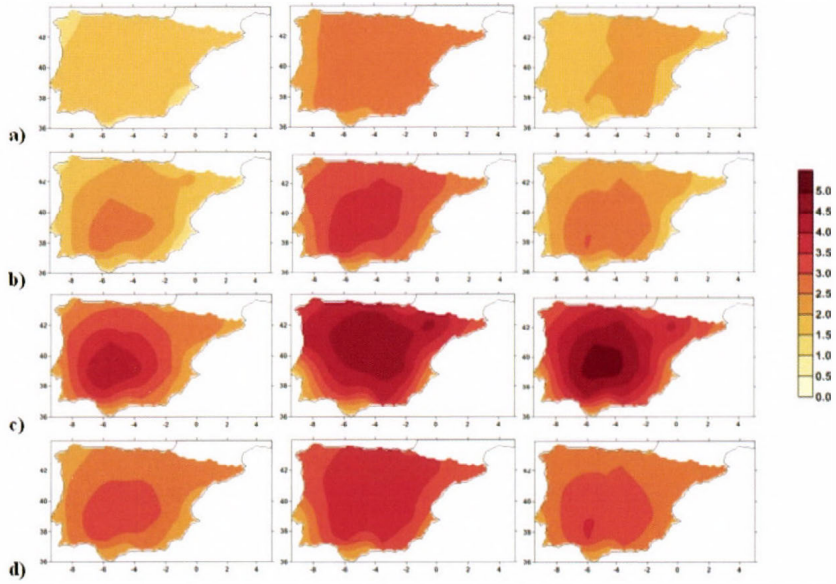

Fig. 15. Seasonal temperature change (in °C) projected for the period 2051–2090 relative to 1961-2000 in the IP for winter (DJF); **a)**, spring (MAM); **b)**, summer (JJA); **c),** and autumn (SON); **d)** using the model outputs derived from DMI-HIRHAM5 driven by ECHAM5-r3 (left), HadRM3 driven by HadCM3 (middle), and KNMI-RACMO2 driven by ECHAM5-r3 (right).

A clear increase in mean seasonal temperature has been projected from all RCM simulations over the whole IP along the year for the periods 2011–2050 and 2051–2090 relative to 1961–2000 (*Figs 14* and *15*). The highest rates were focused in summer and autumn in both time-periods, and the lowest rates were found in winter and spring. DMI-HIRHAM5 and KNMI-RACMO2 outputs showed a 0.5 ºC increase in mean temperatures for winter and spring using the period 2011–2050 with higher rates of 1.5–2 ºC by the 2051–2090 period for the whole IP. Both RCMs produced an increase of 1–1.5 ºC in summer and of 1 ºC in autumn for the period 2011–2050, while they projected an increase of 3–5 ºC in summer and of 2.5–3.5 ºC in autumn by the 2051–2090 period. HadRM3 simulations showed increases of 1.5 ºC for winter, of 1.5–2 ºC for spring and autumn, and of 2–2.5 ºC in summer for the period 2011–2050, while by the 2051–2090 period the increases reached 3 ºC in winter, 3–4 ºC in spring and autumn, and 3.5–4.5 ºC in summer.

All RCM simulations concurred in producing higher rates of change in mean temperatures over the continental IP including the Ebro basin, the Central System, the North and South Plateaus, and the south-western area for all seasons, but especially in summer and autumn. Otherwise, the smallest anomalies were found along the coastline, but especially in the north and north-western coast. In this way, RCM simulations suggested that the continental effect in temperatures will play a major role over the IP along the 21st century reaching extreme temperature increases in the interior area, especially in summer.

5. Discussion and conclusions

In this study, three regional climate models obtained from the EU-Ensembles project have been used to assess projected changes in mean seasonal precipitation and temperature over the whole IP under the A1B climate change scenario for the 21st century. These RCMs were the best-performing models for simulating precipitation and temperature for Europe according to *van der Linden* and *Mitchell* (2009), *Christensen et al.,* (2010), and *Kjellström et al.,* (2010).

The RCM outputs driven by ERA-40 reanalysis have been compared directly with observed MITPS to test the reliability of the simulations. According to the measures tested in this study, KNMI-RACMO2 is the best RCM for simulating precipitation in the IP being consistent with the results obtained by *Simpson* (2011) for UK precipitation, although some problems in summer precipitation should be appreciated. DMI-HIRHAM5 is the best regional climate model for simulation of temperature in the IP. The RCM comparisons with observed data are a necessary, but not sufficient condition to test the accuracy of the models, because current climate change can modify the original basis for a reliable simulation of past climate conditions.

Simulations of seasonal precipitation and temperature from the RCMs driven by GCMs have been used to project the mean expected changes for the periods 2011–2050 and 2051–2090 relative to 1961–2000. Large differences have been detected between RCM outputs coupled with associated GCMs and the observed MITPS. This produces large uncertainties in the results, and these have to be taken into account when assessing model outputs (*Blenkinsop et al.*, 2007; *Sheffield* and *Wood*, 2008; *Rammukainen*, 2010; *Mishra*, 2011; *IPCC*, 2012). Despite, all RCM simulations are projecting a clear decrease of 10% in mean precipitation for spring, summer, and autumn in most IP for the period 2011–2050, it is more evident in the southern than in the northern area. A significant decrease of 20–40% in mean precipitation is expected by the 2051–2090 period for the same seasons, while no clear signal was found in mean winter precipitation from all the RCM outputs and for both time-periods. Model outputs have also shown an increase in mean temperatures between 0.5–1.5 ºC for winter and spring for the period 2011–2050 with higher rates of 1.5–2.5 ºC and 1.5–3 ºC, respectively, by the 2051–2090 period. An increase of between 1–2 ºC was found in summer and 1–1.5 ºC in autumn for the period 2011–2050 being higher by the 2051–2090 period (3–4 ºC in summer and 2.5–3.5 ºC in autumn). Moreover, all RCM simulations concurred by finding higher rates of change in mean temperatures over the continental IP, and the smallest anomalies were found along the coastline, but especially in the north and north-western coasts. In this way, RCM simulations suggested that the continental effect in temperatures will be enhanced and will play a major role in producing extreme temperature increases in the interior area, especially in summer. These findings are consistent with the results obtained by *Gómez-Navarro et al.*, (2010), *Rodríguez-Puebla* and *Nieto,* (2010), and *Jerez et al.*, (2012), and *Jerez* and *Montavez* (2012) for the IP along the 21st century.

The decrease in mean precipitation and the increase in mean temperature projected from the model outputs in the IP could worse current drought conditions for the second half of the 21st century, especially in summer.

Acknowledgements: The authors acknowledge the contribution of *Mª Antonia Valente* and *Ricardo Trigo* from the Climatology and Climate Change Research Group of the Instituto *Dom Luiz*, Lisbon University, to provide us climatic data of Portugal (Porto and Lisboa time series). Original and updated Spanish daily data not included in SDATS and SDAPS were obtained from AEMET servers. Monthly simulated temperature and precipitation data from the three regional climate models driven by ERA-40 reanalysis and driven by GCMs were obtained from the EU-Ensembles project.

References

Aguilar, E., Brunet, M., Saladié, O., Sigró, J., and *López, D.,* 2002: Hacia una aplicación optima del standard normal homogeneity test para la homogeneización de series de temperatura. In (Eds.:*Cuadrat JM, Vicente SM, SAZ MA)* La Información Climática Como Herramienta de Gestión Ambiental. VII Reunión Nacional de Climatología, Grupo de Climatología de la AGE, Universidad de Zaragoza: Zaragoza, 17–33. (In Spain)

Alexandersson, H. and *Moberg, A.*,1997: Homogenisation of Swedish temperature data, Part I: Homogeneity test for linear trends, *Int. J. Climatol. 17*, 25–34.

Barrera-Escoda, A., 2008. Evolución de los extremos hídricos en Catalunya en los últimos 500 años y su modelización regional (Evolution of hydric extremes in Catalonia during the last 500 years and its regional modelling), Ph.D. Thesis, Internal Publication, University of Barcelona, Barcelona, Spain. (In Spain)

Beniston, M., Stephenson, D.B., Christensen, O.B., Ferro, C.A.T., Frei, C., Goyette, S.,Halsnaes, K., Holt, T., Jylha, K., Koffi, B., Palutikof, J., Scholl, R., Semmler, T., and *Woth, K., 2007:* Future extreme events in Eurpoean climate: an exploration of regional climate model projections. *Climatic Change 81*, 71–95.

Bosilovich, M.G., Chen, J., Robertson, F.R., Adler, R.F., 2008: Evaluation of Global Precipitation in Reanalysis. *J. Appl. Meteorol. Climatol. 47*, 2279–2299.

Blenkinsop, S. and Fowler, H.J., 2007: Changes in European drought characteristics projected by the PRUDENCE regional climate models. *Int. J. Climatol. 27*, 1595–1610.

Brunet, M., Jones, P.D., Sigró, J., Saladié, O., Aguilar, E., Moberg, A., Della-Marta, P.M., Lister D., Walther, A., and *López, D.*, 2007: Temporal and spatial temperature variability and change over Spain during 1850-2005. *J. Geophys. Res.-Atmos. 112*:D12117.

Brunet, M., Saladié, O., Jones, P.D., Sigró, J., Aguilar, E., Moberg, A., Walther, A., Lister, D., López, D., and *Almarza, C.*, 2006: The development of a new daily adjusted temperature dataset for Spain (1850–2003). . *Int. J. Climatol. 26*,1777–1802.

Christensen, J.H., Kjellström, E., Giorgi, F., Lenderink, G., and *Rammukainen, M.*, 2010: Weight assignment in regional climate models. *Climate Res. 44*, 179–194.

Christensen, J.H. and *Christensen, O.B.*, 2007: A summary of the PRUDENCE model projections of changes in European climate during this century. *Climatic Change 81*(S1), 7–30.

CLIVAR Assessment, 2010: Clima en España: Pasado presente y futuro. In (Eds. *Pérez F, Boscolo R,*) www.clivar.es (Last visit: 23-02-2013).

Collins, W.J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Hinton, T., Jones, C.D., Liddicont, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Totterdell, I., Woodward, S., Reichler, T., Kin, J., 2010: Evaluation of the HadGEM2 model. *Met Office Hadley Centre. Technical Note no. HCTN 74*, available from Met Office, FitzRoy Road, Exeter, UK. EX1 3PB.

Dai, A., 2011. Drought under global warming: A review. *Wiley Interdisciplinary Reviews:Climate Change 2*, 45–65.

Dai, A., 2012. Increasing drought under global warming in observations and models. *Nature Climate Change 3, 53–58.*

ECMWF 2004. ERA-40: ECMWF 45-year reanalysis o the global atmosphere and surface conditions 1957–2002. *ECMWF Newsletter Nº. 101.* Retrieved 5th November, 2010; http://www.mad.zmaw.de/uploads/media/e40Overwiew.pdf.

Gomez-Navarro, J.J., Montavez, J.P., Jimenez-Guerrero, P., Jerez, S., Garcia-Valero, J.A., Gonzalez-Rouco, J.F., 2010: Warming patterns in regional climate change projections over the Iberian Peninsula. *Meteorol. Z. 19*, 275–285.

IPCC climate change 2012: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. *(Eds. Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley),* Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582.

IPCC climate change 2007: the physical science basis. In: (*Eds .Salomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miler HL,)* Contribution of Working Group I to the Fourth Assessment Report of the International Panel on Climate Change Program. Cambridge, UK/New York, USA: Cambridge University Press; 2007, 996.

Jerez, S., Montavez, J.P., Gomez-Navarro, J.J., Jimenez, P.A., Jimenez-Guerrero, P., Lorente-Plazas, R., Gonzalez-Rouco, J.F., 2012: The role of the land-surface model for climate change projections over the Iberian Peninsula. *J. Geophys. Res. 117*, D01,109.

Jerez, S. and *Montavez, J.P.*, 2012: A multi-physics ensemble of regional climate change projections over the Iberian Peninsula. *Clim. Dynam.* DOI: 10.1007/s00382-012-1551-5.

Kjellström, E., Boberg, F., Castro, M., Christensen, JH., Nikulin, G., and *Sanchez, E.,* 2010. Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. *Climate Res. 44*, 121–134.

van der Linden and *Mitchell, J.F.B.,* 2009: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK.

Mariotti, A., Zeng, N., Yoon, J.H., Artale, V., Navarra, A., Alpert, P., and Li, L.Z.X., 2008: Mediterranean water cycle changes: transition to drier 21st century conditions in observations and CMIP3 simulations. *Environ. Res. Letters, 3,* 044001.

Mishra, A.K. and *Singh, V.P.,* 2011: Drought modeling-A review. *J. Hydrol. 403*, 157–175.

Rammukainen, M., 2010: State-of-the-art with regional climate models. *Wiley Interdisciplinary Reviews: Climate Change 1*, 82–96.

Rodríguez-Puebla, C. and *Nieto, S.,* 2010. Trends of precipitation over Iberian Peninsula and the North Atlantic Oscillation under climate change conditions. *Int. J. Climatol. 30*, 1807–1815.

Sánchez, E., Gaertner, M.A., Gallardo, C., 2009: Regionalización diaria de la precipitación diaria sobre la Península Ibérica: análisis de la resolución espacial en la descripción del clima actual y clima futuro. *Física de la Tierra 21*, 207–218. (In Spain)

Sheffield, J. and *Wood, E.F.,* 2008. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Clim. Dynam. 31*, 79–105.

Simpson, I., 2011. PhD Dissertation. Precipitation variability across the UK: Observations and model simulations. Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich NR47TJ, UK.

Sousa P., Trigo, R.M., Aizpurua, P., Nieto, R., Gimeno, L., Garcia-Herrera, R., 2011: Trends and extremes of drought indices throughout the 20th century in the Mediterranean, NHESS, Special Issue "*Understanding dynamics and current developments of climate extremes in the Mediterranean region", 11*, 33–51.

Van Meijgaard, E., van Ulft, L.H., van de Berg, W.J., Bosvald F., van den Hurk, B., Lenderink, G., and *Siebesma A.,* 2008*:* The KNMI regional atmospheric climate model RACMO versión 2.1. *KNMI Techn. Report TR-302,* available at: www.knmi.nl/publications/fulltexts/tr302_racmo2v1.pdf

Vicente-Serrano, S.M., Beguería, S, López-Moreno, J.I., 2011: Comment on "Characteristics and trends in various forms of the Palmer Drought Severity Index (PDSI) during 1900-2008" by Aiguo Dai, *J. Geophys. Res., 116*, D19112.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 1, January – March, 2015, pp. 23–39

Evaluation and gap filling of soil NO flux dataset measured at a Hungarian semi-arid grassland

Dóra Hidy[1], **László Horváth**[*1], and **Tamás Weidinger**[2]

[1]*MTA-SZIE Plant Ecology Research Group, Szent István University, Páter K. u. 1, 2103 Gödöllő, Hungary*

[2]*Department of Meteorology, Eötvös Loránd University, Pázmány Péter sétány 1/A, 1117 Budapest, Hungary*

[*]*Corresponding author E-mail: horvath.laszlo.dr@gmail.com*

(Manuscript received in final form February 8, 2015)

Abstract—Nitric oxide soil emission flux was measured by 2–2 parallel manual and auto dynamic chambers on hourly basis above a Hungarian semi-arid, sandy grassland between August 2012 and January 2014. The measured datasets covered 43–85% of time period depending on chambers. We applied a gap filling method based on multivariable analysis (Sigma Plot) combined with maximum likelihood method. Trend of gap filled dataset shows large peaks mostly in summer and early fall. When soil parameters are far from the optimum (dry, warm conditions), the fluxes are negligible. Application of manual chambers closed for longer period results in substantial positive bias in flux estimation compared to auto chambers as a consequence of measurement setup, different temperature, and drier soil conditions below the chamber. Mean fluxes applying permanently closed dynamic chambers are approximately three times higher compared to auto chambers: 0.176±0.489 nmol $m^{-2}s^{-1}$ and 0.058±0.130 nmol $m^{-2}s^{-1}$, respectively.

Key-words: nitric oxide, soil emission, gap filling, nitrification, grassland, dynamic chamber

1. Introduction

As it is well-known, one of the most important natural sources of atmospheric nitric oxide (NO) is the soil nitrification-denitrification process. It has been recognized earlier that soil flux of NO is similar in magnitude to fossil fuel emission of NO_x (*Davidson* and *Kingerlee*, 1997). For this reason, NO plays key role both in biosphere–atmosphere N-balance and in biogeochemical cycle of nitrogen.

Rate of soil NO flux strongly depends on soil temperature and moisture (*Smith et al.*, 1998) and on soil aeration, inorganic N content, and pH (*Bouwman*, 1996; *Cárdenas et al.*, 1993; *Pilegaard*, 2013).

There have been several research projects aiming on one hand to establish the magnitude of emission rate of nitrogen oxides from soils (e.g., NOFRETETE, *Kesik et al.*, 2005; *Pilegaard et al.*, 2006), and on the other hand to involve the soil NO emission in the N-inventory (e.g., NitroEurope, *Skiba et al.*, 2009). An Integrated Project ÉCLAIRE (http://www.eclaire-fp7.eu/) started in 2011, among others to study the effect of climate change on air pollution impacts. One of the main measurement tasks of this project was the continuous monitoring of soil NO emission at different types of land (forest, arable, grass). Among the tree European grass stations, Bugacpuszta (Hungary) was selected to monitor and report soil NO fluxes continuously for 17 months on hourly basis.

We applied parallel both so-called manual (permanently closed) and auto (closed only during short measurement period) chambers for continuous measurements. As it was established earlier (*Yao et al.*, 2009), the use of manual chamber for measuring soil fluxes has some disadvantages. One of the main disadvantages is caused by the permanently closed status of chambers preventing the surface from precipitation. On the other hand, the top of chamber is exposed to solar radiation continuously heating the chamber box inside, compared to auto chambers which are open out of measurement cycle. Lower moisture and higher temperature conditions inside the permanently closed chambers may generate systematic bias in flux calculation.

As a consequence of the malfunction and failure of sampling and monitoring equipments during our measurement period, the data covers only the 43–85% of the full time of measurement campaign for different chambers.

The aim of this paper is firstly to apply a statistical gap filling method to complete the dataset, and secondly to give a semi-quantitative estimation on the positive bias in soil NO fluxes caused by applying manual chambers.

2. Methodology

2.1. Measurement of soil NO fluxes

Measurements were carried out in Bugacpuszta, Hungary, above a grassland between August 2012 and January 2014. The climate is semi-arid temperate continental, the mean annual temperature is 10.7 °C, and the average yearly precipitation is around 550 mm. The region has sandy soil with high sand (79%) and low clay (13%) contents in the upper 10 cm soil layer. More detailed information of location and characteristics are described in *Horváth et al.* (2010); *Machon et al.* (2010; 2011; 2015). Two-two parallel manual (Chamber 1 and Chamber 2; V = 2 L, h = 5 cm) and automatic (Chamber 3 and Chamber 4; V = 6.8 L, h = 10 cm) dynamic chambers were applied. Automatic chambers (Ricambi, Milan) were settled at fix positions and were closed only for 10 minutes every hour during sampling allowing the surface to be exposed to solar radiation and precipitation out of sampling time. Two manual chambers, home made by pale grey plastic, were permanently closed. The two chambers were re-settled during bi-weekly station maintenances onto other collars among the 6 fixed ones in turn. All of four chambers were sampled for 10 minutes at a flow rate of 2 L min^{-1} in sequence every hour all together for 40 minutes; in the remaining 20 minutes, concentration gradients were measured by a mast at different heights. Soil temperature and moisture were measured a few meters apart from the chambers by 105T thermocouple probes and CS616 water content reflectometers, at –5; –30 cm and –3; –30 cm depths, respectively. A computer controlled valve system was switched the different channels in turn. The output concentrations of nitric oxide and ozone were measured by HORIBA APNA-350E and APOA-350E gas monitors through teflon tubing. Input concentrations of these gases were estimated from gradient concentration measurements at 0.5 m height, at the beginning of the one-hour long measurement cycle.

NO flux was calculated according to *Meixner et al.* (1997). Chemical correction of rapid reaction of NO with ozone ($NO+O_3 \rightarrow NO_2+O_2$) was taken into account. Under steady-state conditions, the mass balance equation for NO can be written as follows (the photolysis rate of NO_2 inside the dark chambers was estimated to be zero):

$$F_f + F_m + F_{bl} + F_{gp} = 0, \tag{1}$$

where F_f is the soil flux, F_m is the difference between fluxes entering and leaving the chamber, F_{bl} is the term for the wall effect which was negligible because of the relatively short residence time of the gas mixture in the chamber, and F_{gp} is the loss of NO due to the chemical reaction with ozone. For detailed description of flux calculations refer to *Horváth et al.* (2006) and *Machon et al.* (2015).

2.2. *Statistical gap filling method*

As soil NO flux depends on soil moisture, temperature, and soil organic nitrogen content, we have taken into account the variation of soil flux as a function of these parameters for gap filling. In the lack of regular observations we supposed that chemical characteristics of soil did not changed significantly during the observation period.

There are different gap filling methods used in micrometeorological flux measurements (*Papale*, 2012) i) empirical (e.g., look up tables), ii) interpolation (e.g., mean diurnal variation), iii) artificial neutral networks, iv) non-linear regressions, and v) process oriented models. Gap filling methodology has been applied for the long term eddy covariance dataset in most cases. A statistical–regression methodology for gap filling of the long term soil respiration measurements was constructed by *Gomez-Casanovas et al.* (2013).

Effects of soil temperature and moisture for soil NO fluxes are well investigated and described by Gaussian distributions (*Luo et al.*, 2013; *Pilegaard*, 2013). Based on these investigations and on the former analysis of soil fluxes in Bugacpuszta (*Machon et al.*, 2011), a nonlinear regression gap filling method was selected which consisted of different steps.

Firstly we analyzed the dependence of soil flux in the function of soil physical parameters. We have taken into account four parameters, namely the soil moisture measured at –3 and –30 cm and soil temperature measured at –5 and –30 cm depths as SWC1, SWC2, TS1, and TS2, respectively (*Fig. 1*). During the measurement period, a total of 4686 parallel measurements were taken when all of the 4 chambers were together in operation. The dependence of fluxes on physical parameters was tested only for these measurements ensuring the homogeneity. We supposed that shape of functions is the same for all the measuring plots (chambers).

According to the shape of these functions and on the basis of earlier observations at the same site (*Machon et al.*, 2011), we supposed a maximum shape, exponential functions for moisture (x_{SWC}) and temperature (x_{TS}) using the assumption of Gaussian distributions. Although the flux dependence on soil temperature is generally exponential, in our case the temperature often exceeded the 20 °C resulting in lower bacterial activity caused by the heat stress or by the extreme low humidity at higher temperature regimes. As the first step, Eq. (2) was used for the estimation of missing soil flux rates (F_f):

$$F_f = a_{SWC} \cdot \exp\left[-0.5 \cdot \left(\frac{x_{SWC} - x0_{SWC}}{b_{SWC}}\right)^2\right] + a_{TS} \cdot \exp\left[-0.5 \cdot \left(\frac{x_{TS} - x0_{TS}}{b_{TS}}\right)^2\right]. \quad (2)$$

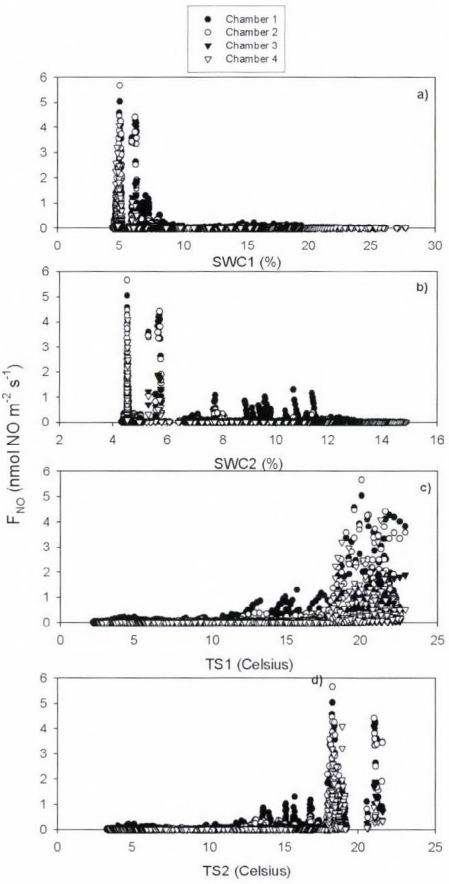

Fig. 1. NO fluxes measured by the four chambers in function of volumetric soil water content SWC1 (a), SWC2 (b) at –3 and –30 cm and soil temperature TS1 (c), TS2 (d) at –5 cm and –30 cm depths, respectively.

3. Results

3.1. Gap filling of data series

We have calculated the dependence of NO flux in 4 different combinations of SWC and TS by the SigmaPlot 8.0 (Systat Software Inc., Chicago, USA) graphing and data analyses. (The Sigma Plot curve fitter uses the Marquardt-Levenberg algorithm). The best fit was observed among soil fluxes and SWC1-TS1 (*Fig. 2*) among hourly data on the days (n=4686) when no measurement was missing (all chambers were in operation). Measured flux data

ranges within 0–6 nmol $m^{-2} s^{-1}$. The relationship is significant at the probability level of $p < 0.0001$ (except of two cases), and the calculated correlation coefficients are: $R^2_{Ch1} = 0.435$; $R^2_{Ch2} = 0.436$; $R^2_{Ch3} = 0.393$; $R^2_{Ch4} = 0.323$. The standard errors (standard deviation of differences between the measured and estimated fluxes) of estimation are: $SD_{Ch1} = 0.233$; $SD_{Ch2} = 0.233$; $SD_{Ch3} = 0.079$; $SD_{Ch4} = 0.139$ nmol $m^{-2} s^{-1}$.

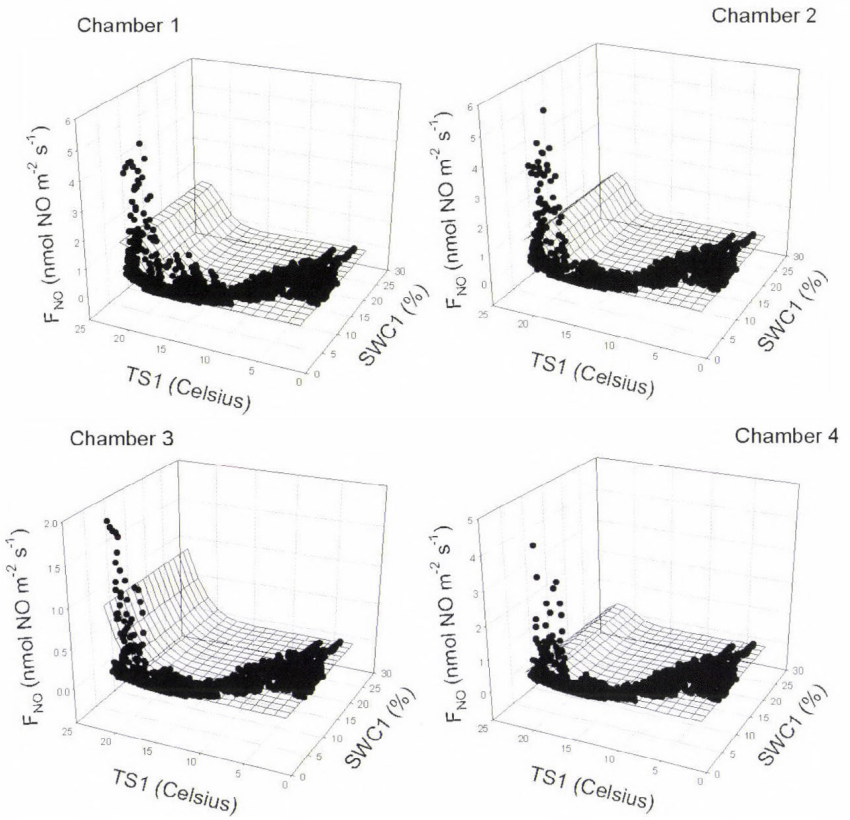

Fig. 2. Dependence of NO flux on soil properties (SWC at –3 cm depth, TS at –5 cm depth) measured by the different chambers. The fitted surface was calculated based on Eq. (2), the parameter estimation was created by Sigma Plot.

Further on, we have estimated the parameters of Eq. (2) by maximum likelihood method. There are various measures to express the distance between model and observation numerically (e.g., average, median, maximum, and normalized). It is difficult to clearly judge the significance of the different

quantitative measures, the choice of the misfit depends on the nature of the reference data (*Janssen* and *Heuberger*, 1995).

Hence, the NO flux estimation has two functions with 3–3 parameters for each chamber. In the next step, the parameter values were randomized for all chambers (Chamber 1–Chamber 4) to find the best combination of them. The measure of the goodness-of-fit is a special likelihood (Appendix Eq. A5.) taking simultaneously into account the correlation between the measured and modeled data, the average error, and the difference of the sum of measured and estimated data (see the equations in detail in the Appendix) (*Janssen* and *Heuberger*, 1995). The parameter values belonging to the maximum of the likelihood values are considered as the optimum parameter set.

After estimation for each chamber, the model was run with the optimum parameters compiled in *Table 1*. As the result of calibration, the correlation remained the same among measured and calculated data with parallel decrease of the error of estimation and the difference between measured and simulated sum (*Table 2*).

Table 1. Optimized parameters of Eq. (2)

	a_{SWC}	b_{SWC}	$x0_{SWC}$	a_{TS}	b_{TS}	$x0_{TS}$
Chamber 1	0.003	0.751	26.409	2.689	1.892	22.298
Chamber 2	0.050	0.227	48.252	2.060	1.571	21.399
Chamber 3	–1.524	–0.268	50.073	25.542	4.805	33.980
Chamber 4	2.134	0.073	3.068	1.044	1.813	21.505

Table 2. Coefficient of determination (R^2), normalized error (NE, nmol $m^{-2} s^{-1}$), and the difference between the measured and simulated sum (SE, nmol $m^{-2} s^{-1}$), before (BC, Method 1) and after (AC, Method 2) calibration (n =4686)

	Chamber 1			Chamber 2			Chamber 3			Chamber 4		
	R^2	NE	SE	R^2	NE	SE	R^2	NE	SE	R^2	NE	SE
BC	0.44	1.36	–32.8	0.44	1.27	–28.8	0.39	1.52	–55.1	0.32	1.6	–20.1
AC	0.45	1.32	–1.82	0.45	1.26	0.74	0.37	1.58	6.55	0.29	1.1	10.9

The next step of gap filling procedure was the estimation of goodness of method. We analyzed the distribution of the data gaps. Each data series (Chambers 1–4) has 12,744 rows, with 15% to 57% of data lack depending on the chamber. We found that the most frequent length of the data gaps was 1, and the maximum was 689 in hour scale (*Fig. 3*).

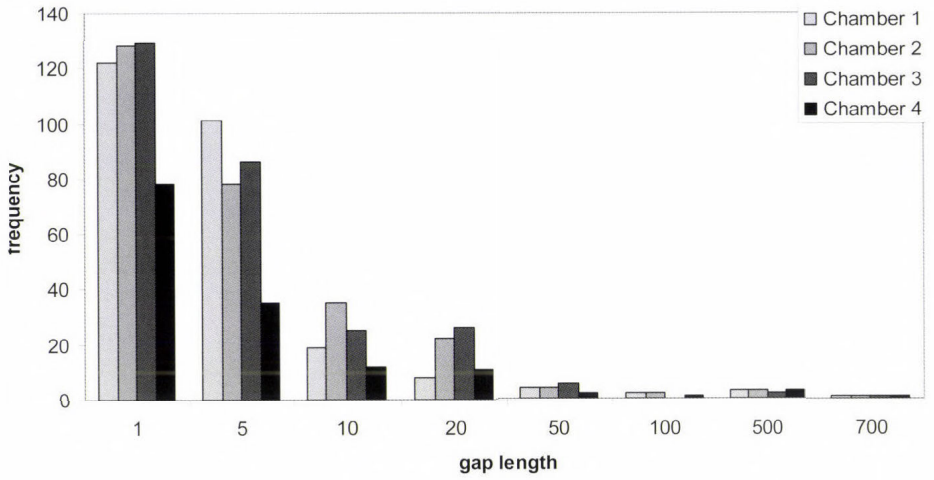

Fig. 3. Distribution of the data gaps (number of missing hourly data) regarding to the different chambers.

In order to examine the efficiency of gap filling methods, we created different numbers of data lacks in the measured data series by random number generator. *Table 3* contains the length and number of artificial gaps selected according to the original distribution of the data gaps (*Fig. 3*). The random gap creation was repeated 1000 times. The lacks were filled both by the simple method (Method 1: linear interpolation between the last data before the given gap and the first data after the given gap) and by the method described above (Method 2: using NO flux estimation based on SWC and TS data). The measured (NO flux data set without artificial gaps) and the estimated (artificial gaps are filled with estimated data) NO flux data were compared using likelihood values (function of the difference between the measured and estimated NO data defined by Eq. A5. in the Appendix) during random gap creation (1,000 random gap – 1,000 likelihood value). Hence, by the comparison of the two data series generated by the two methods (difference of the average likelihood), we found that in case of manual chambers, Method 1 gives better estimation (the difference is positive) if gap length is less than 10 hourly data and in case of automatic chambers, Method 1 is the better if gap length is less than 30. We can verify that we can use Method 2 for manual chambers in cases when gap length is longer than 10 and for automatic chambers in cases when gap length is longer than 30 (in other cases, Method 1 was used to fill the gaps) (*Fig. 4*).

Table 3. Length and number of artificial data gaps

name	length	number
G1	1	100
G5	5	100
G10	10	20
G20	20	10
G20	30	5
G50	50	5
G100	100	1
G500	500	1
G700	1000	1

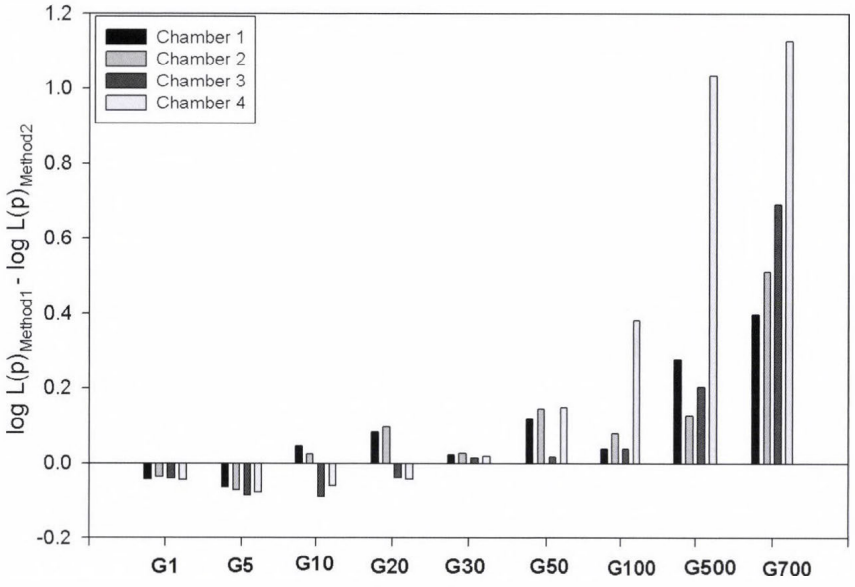

Fig. 4. The difference between the likelihood values using Method 1 and Method 2 during random gap creation in function of the gap length (G1: 1 data is missing, G5: five data are missing, etc.) for different chambers.

3.2. Evaluation of data

For evaluation, one dataset among 4 chambers was selected supposing that trend of NO flux is similar for all. The gap filled data series for Chamber 4 can be seen in *Fig. 5*. The trend of NO fluxes follows well the variation of soil wetness and temperature as it is expected. The optimum conditions for nitrification and

NO emission are in the dry and warm ranges. Two optimum periods occurred during summer/early fall during years of 2012 and 2013, when emission peaks appeared. In any other time period at either low temperature (lower than 10 °C) or high soil water content (higher than 10%), the NO production/emission are suppressed; the soil fluxes are negligible in magnitude compared to summer and early fall rates characterized by optimum soil conditions.

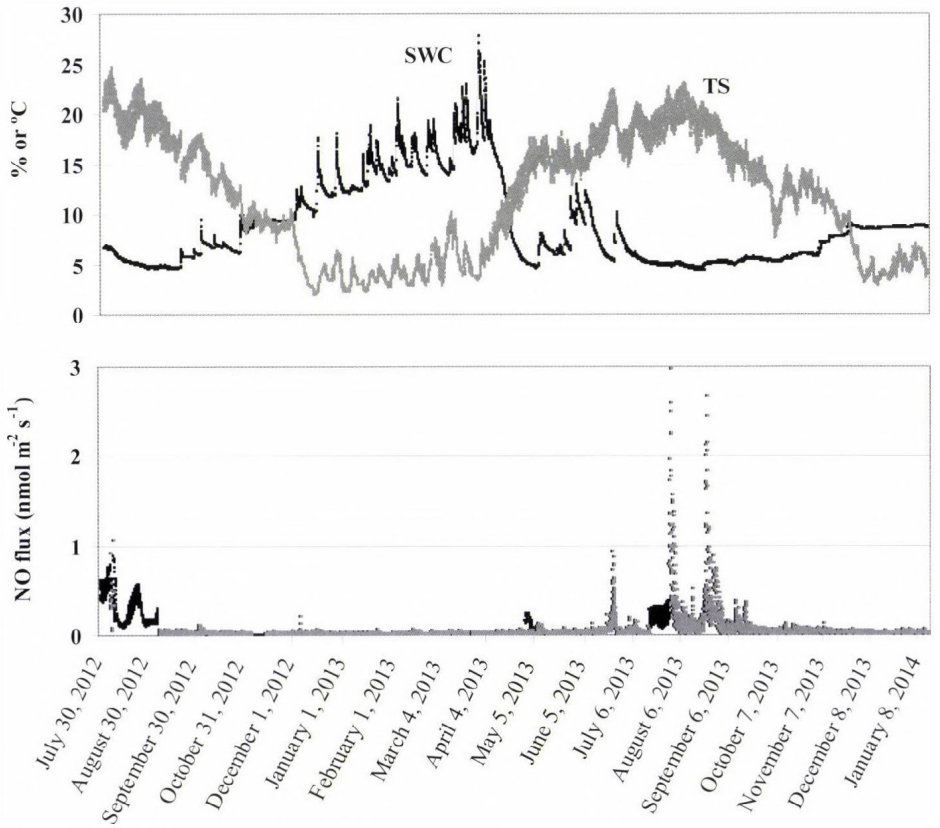

Fig. 5. Time course of soil properties, and measured (grey), and gap filled (black) NO fluxes for Chamber 4.

The statistical parameters of data series can be seen in *Table 4.* Gap filled flux data sets were compared to modeled ones for the whole period and for year 2013, separately. The agreement is acceptable for Chambers 1–3; the largest deviation

from the mean (less than 20%) appears for Chamber 4. Coefficients of variation (CV) for observed and calculated fluxes agree in magnitude for all chambers.

Table 4. Statistical parameters (mean, CV*) of measured and gap filled datasets for soil NO flux (nmol $m^{-2}s^{-1}$)

	Chamber 1	Chamber 2	Chamber 3	Chamber 4
Modeled, all				
mean	0.168	0.183	0.062	0.065
CV	3.16	2.59	2.58	2.46
n	12 747	12747	12 747	12 747
Gap-filled, all				
mean	0.171	0.181	0.061	0.055
CV	3.04	2.54	2.41	2.05
n	12 747	12 747	12 747	12 747
Modeled, 2013				
mean	0.086	0.122	0.042	0.058
CV	2.88	2.21	2.65	2.82
n	8 760	8 760	8 760	8 760
Gap-filled, 2013				
mean	0.094	0.113	0.041	0.044
CV	2.11	1.93	1.94	1.92
n	8 760	8 760	8 760	8 760

CV* (coefficient of variation): ratio of the sample standard deviation to the sample mean

3.3. Estimation of bias by using manual chambers

Mean of hourly soil NO fluxes were calculated separately for manual and auto chambers for cases when all of the 4 chambers were parallel in operation (n = 4686). The bulk daily course of ratio of fluxes measured by manual and auto chambers can be seen in *Fig. 6.* At night when solar radiation was zero, the average positive bias in fluxes measured by manual chambers (independently from the season) can be characterized by a factor of 1.6 as a consequence of drier soil conditions below the chamber. When solar radiation reached its maximum around the noon hours, the factor has increased up to 3.2. Mean flux values in *Table 4* also demonstrate a huge (around a factor of 3) increase in fluxes by applying permanently closed chambers. Because the magnitude of the

bias depends on many factors (climate, material and dimensions of chamber etc.), deeper conclusion can not be drown out of this semi-quantitative estimation.

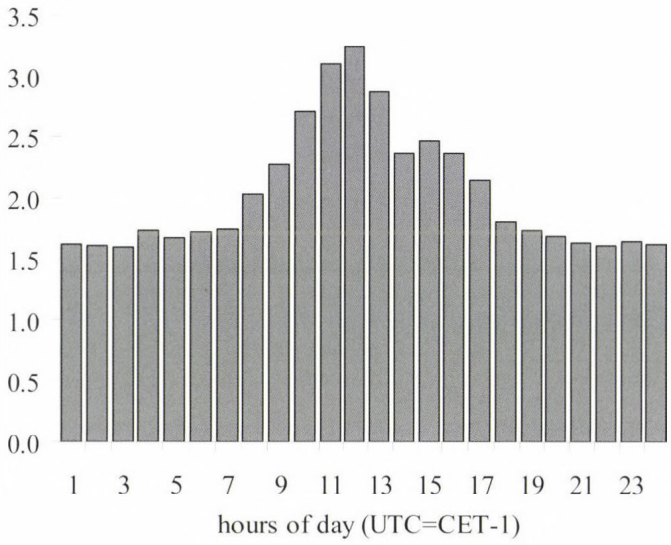

Fig. 6. Daily course of bias as ratio of measured soil NO flux by manual chambers compared to auto chambers (n=4686).

4. Conclusion

Soil nitric oxide flux over sandy grassland strongly depends on soil moisture and temperature. Using the significantly correlated functions among fluxes, soil moisture, and temperature, the missing fluxes can be predicted. Analyzing the whole dataset we can establish that significant fluxes were measured when temperature and soil wetness were near to the optimum rate. In any cases, the fluxes were practically negligible. Expecting a drier and warmer climate in our region, reduction in soil NO emission is expected in the future.

The application of manual chambers (closed for longer period) for soil flux measurement may cause significant positive bias especially through the heating effect of solar radiation. The use of white chambers may reduce this effect, but the lid of chambers is an obstruction for precipitation that can not be prevented.

Acknowledgements—Authors acknowledge the financial support of Animal Change (FP7 266018), ÉCLAIRE (FP7 282910), NitroEurope (FP6) EU projects, MTA PD 450012 project, and PIAC 13-1-2013-0141 research and development project. Special thanks to Gyula Pávó, Zoltán Istenes, and Attila Eredics for their technical help.

References

Bouwman, A.F., 1996: Direct emission of nitrous oxide from agricultural soils. *Nutr. Cycl. Agroecosys. 46*, 53–70.

Cárdenas, L., Rondón, A., Johansson, Ch., and *Sanhueza, E.*, 1993: Effects of soil moisture, temperature, and inorganic nitrogen on nitric oxide emissions from acidic tropical savannah soils. *J. Geophys. Res.: Atmos. 98*, 14783–14790.

Davidson, E.A. and *Kingerlee, W.*, 1997: A global inventory of nitric oxide emissions from soil. *Nutr. Cycl. Agroecosys. 48*, 37–50.

Gomez-Casanovas, N., Anderson-Teixeira K., Zeri, M., Bernacchi, C.J., and *DeLucia, E.H.*, 2013: Gap filling strategies and error in estimating annual soil respiration. *Glob. Change Biol. 19*, 1941–1952.

Horváth, L., Führer, E., and *Lajtha, K.*, 2006: Nitric oxide and nitrous oxide emission from Hungarian forest soils; linked with atmospheric N-deposition. *Atmos. Environ. 40,* 7786–7795.

Horváth, L., Grosz, B., Machon, A., Tuba, Z., Nagy, Z., Czóbel, Sz., Balogh, J., Péli, E., Fóti, Sz., Weidinger, T., Pintér, K., and *Führer, E.*, 2010: Estimation of nitrous oxide emission from Hungarian semi-arid sandy and loess grasslands, effect of soil parameters, grazing, irrigation and application of fertilizer. *Agric., Ecosys. Environ.139*, 255–263.

Janssen, P.H.M., and *Heuberger, P.S.C.*, 1995: Calibration of process oriented models. *Ecol. Model. 83,* 55–66.

Kesik, M., Ambus P., Baritz, R., Brüggemann, N., Butterbach-Bahl, K., Damm, M., Duyzer, J., Horváth, I., Kiese, R., Kitzler, B., Leip, A., Li, C., Pihlatie, M., Pilegaard, K., Seufert, G., Simpson, D., Skiba, U., Smiatek, G., Vesala, T., and *Zechmeister-Boltenstern, S.*, 2005: Inventory of N_2O and NO emissions from European forest soils. *Biogeosciences 2,* 353–375.

Luo, G.J., Kiese, R., Wolf, B., and *Butterbach-Bahl, K.*, 2013: Effects of soil temperature and moisture on methane uptakes and nitrous oxide emissions across three different ecosystem types. *Biogeosciences 10*, 3205–3219.

Machon, A., Horváth, L., Weidinger, T., Grosz, B., Pintér, K., Tuba, Z., and *Führer, E.*, 2010: Estimation of net nitrogen flux between the atmosphere and a semi-natural grassland ecosystem in Hungary. *Eur. J. Soil Sci. 61*, 631–639.

Machon, A., Horváth, L., Weidinger, T., Pintér, K., Grosz, B., Nagy, Z., and *Führer, E.*, 2011: Weather induced variability of N-exchange between the atmosphere and a grassland in the Hungarian Great Plain. *Időjárás 115*, 219–232.

Machon, A., Horváth, L., Weidinger, T., Grosz, B., Móring, A., and *Führer, E.*, 2015: Measurement and modeling of N-balance between the atmosphere and the biosphere over a grazed grassland (Bugacpuszta) in Hungary. *Water Air Soil Poll. 226*:27.

Meixner, F.X., Fickinger, Th., Marufu, L., Serca, D., Nathaus, F.J., Makina, E., Mukurumbira, L., and *Andreae, M.O.*, 1997: Preliminary results on nitric oxide emission from a southern African savanna ecosystem. *Nutr. Cycl. Agroecosys. 48*, 123–138.

Papale, D., 2012: Data gap filling. In (eds. *Aubinet, M., Vesala, T., and Papale, D.*), Eddy covariance. Springer, Dordrecht, Heidelberg, London, New York.159–172.

Pilegaard K., 2013: Processes regulating nitric oxide emissions from soils. *Phil. Trans. R. Soc. B 368*, 20130126.

Pilegaard, K., Skiba, U., Ambus, P., Beier, C., Brüggemann, N., Butterbach-Bahl, K., Dick, J., Dorsey, J., Duyzer, J., Gallagher, M., Gasche, R., Horvath, L., Kitzler, B., Leip, A., Pihlatie, M.K., Rosenkranz, P., Seufert, G., Vesala, T., Westrate, H., and *Zechmeister-Boltenstern, S.*, 2006: Factors controlling regional differences in forest soil emission of nitrogen oxides (NO and N_2O). *Biogeosciences 3*, 651–661.

Skiba, U., Jones, S.K., Drewer, J., Tang, Y.S., van Dijk, N., Helfter, C., Nemitz, E., Twigg, M., Famulari, D., Owen, S., Philatie, M., Vesala, T., Larsen, K.S., Carter, M.S., Ambus, P., Ibrom, A., Beier, C., Hensen, A., Frumau, A., Brüggemann, N., Gasche, R., Neftel, A., Spirig, C., Horvath, L., Freibauer, A., Cellier, P., Laville, P., Loubet, B., Magliulo, E., Bertolini, T., Seufert, G., Andersson, M., Manca, G., Laurila, T., Aurela, M., Zechmeister-Boltenstern, S., Kitzler, B., Schauffler, G., Siemens, J., Kindler, R., Flechard, C., Sutton, M.A., Erisman, J.W., Cape, J.N., and *Butterbach-Bahl, K.*, 2009: Biosphere atmosphere exchange of reactive nitrogen

and greenhouse gases at the NitroEurope core flux measurement sites: Measurement strategy and first data sets. *Agr. Ecosys. Environ.133*, 139–149.

Smith, K.A., Thomson, P.E., Clayton, H., McTaggart, I.P., and Conen, F., 1998: Effects of temperature, water content and nitrogen fertilisation on emissions of nitrous oxide by soils. *Atmos. Environ. 32*, 3301–3309.

Yao, Zh., Zheng, X., Xie,B., Liu,Ch., Mei,B., Dong, H., Butterbach-Bahl, K., and *Zhu, J.*, 2009: Comparison of manual and automated chambers for field measurements of N_2O, CH_4, CO_2 fluxes from cultivated land. *Atmos. Environ. 43*, 1888–1896.

Appendix

A1: Basic definitions

Error (E): difference between measured and simulated data on a given hour (i) using a given parameter set (p)

$$E_i = m_i(p) - d_i^{obs}$$

Normalized bias (NB): normalized difference between the sum of model predictions and observed values

$$NB = \frac{\overline{M} - \overline{O}}{\overline{O}}$$

Modeling efficiency: a measure used to assess the predictive power of models (definition is identical to coefficient of determination (R^2) in case of linear regression)

$$ME = 1 - \frac{\sum_i (E_i)^2}{\sum_i \left(d_i^{obs} - \overline{d_i^{obs}}\right)^2}$$

Log-likelihood: it is more convenient to calculate with the natural logarithm of the likelihood

$$\log L(p) = \ln[L(p)]$$

A2. Likelihood and log-likelihood from Janssen median error misfit

$$L(p)_{average} = \exp\left[-\left(\frac{1}{N}\sum_{i=1}^{N} NE_i\right)\right]$$

$$\log L(p)_{average} = -\left(\frac{1}{N}\sum_{i=1}^{N} NE_i\right)$$

A3. Likelihood and log-likelihood from Janssen modeling efficiency error misfit

$$L(p)_{efficiency} = \exp[-(ME)^{-1}]$$
$$\log L(p)_{efficiency} = -(ME)^{-1}$$

A4. Likelihood and log-likelihood from Janssen bias misfit

$$L(p)_{bias} = \exp[-NB]$$
$$\log L(p)_{bias} = -NB$$

A5. Combined Likelihood and log-likelihood form

$$L(p)_{combined} = L(p)_{average} \cdot L(p)_{bias} \cdot L(p)_{efficiency}$$
$$\log L(p)_{combined} = -\left(\frac{1}{N}\sum_{i=1}^{N} NE_i\right) + (-ME) + (-NB)$$

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 1, January – March, 2015, pp. 39–51

Radar-based investigation of long-lived thunderstorms in the Carpathian Basin

Ákos Horváth[1]*, **András Tamás Seres**[2], and **Péter Németh**[3]

[1]*Hungarian Meteorological Service,*
Vitorlás u. 17, H-8600 Siófok, Hungary,
E-mail: horvath.a@met.hu

[2]*Hungarian Defence Forces Geoinformation Service,*
Szilágyi E. fasor 7–9, H-1024 Budapest, Hungary,
E-mail: seres.andrastamas@upcmail.hu

[3]*Hungarian Meteorological Service,*
Gilice tér , H-1024 Budapest, Hungary,
E-mail: nemeth.p@met.hu

**Corresponding author*

(Manuscript received in final form April 8, 2014)

Abstract – This study describes a weather-radar-based investigation of long-lived thunderstorms in Hungary in the period of 2004–2012. An objective method was developed for identifying and tracking convective cells. The cells were represented by so-called thunderstorm ellipses. In this research, intensive objects were classified into 3 categories such as severe, highly severe, and extremely severe thunderstorm ellipses. The categories were defined by radar reflectivity thresholds 45 dBZ, 50 dBZ, and 55 dBZ. Only those cells were involved in the investigation whose lifetime extended more than 1 hour. In the 9-year period, 2625 severe, 597 highly severe, and 45 extremely severe long-lived thunderstorm ellipses were found. Stronger cells moved faster and at most intensive cells, right-turning movement was more frequent. Many of these long-lived, strong objects could be supercells. The applied methods and results can be used for severe weather forecast and nowcasting in the Carpathian Basin.

Key-words: severe thunderstorm, weather radar, climatology, tracking, supercell, Hungary

1. Introduction and background

Severe thunderstorms and associated phenomena (stormy wind gusts, hailstorms, heavy rainfalls, tornadoes, intensive lightning) often occur in Hungary, mainly in late spring and summer (*Horváth* and *Geresdi,* 2003). These convective storms cause major damages mainly in the electric and transportation networks by wind gusts, in the agriculture by hail storms, and in other branches of the infrastructure by torrential rain and flash floods. A thunderstorm, developed into supercell phase, caused fatal accidents during the Constitution Day firework in Budapest in 2006 (*Horváth et al., 2007*). These events motivate the investigation to overview severe convective cells using 9-year period of observations by radar network of the Hungarian Meteorological Service. In practice, radar reflectivity data are available in that period, what explains why reflectivity based identification and tracking methods were applied for recognizing and tracking intensive thunderstorms.

Radar-based tracking of thunderstorms have been investigated in the United States since the beginning of the 1950's. *Battan* (1952) found that the storm-tops of "longer-lived" (> 20 min) cells were higher than the shorter-lived (≈10 min) ones. *Browning* (1964) defined *the supercell* term for the most organized, most severe, and longest-lived form of isolated, deep moist convection. In 1966, *Wilson* analyzed the relation of thunderstorm size and intensity (*Wilson,* 1966). In the 1970s, the development of remote sensing technique and the surface observations allowed to update the definitions of supercells (*Wilhelmson* and *Klemp*, 1978; *Lemon* and *Doswell*, 1979). Presently, the most accepted theory about supercell thunderstorms was made by *Klemp* (1987). *Henry* (1993) and *MacKeen et al.* (1999) analyzed the relationship between reflectivity–derived storm characteristics and storm longevity. In the last two decades, some radar-based algorithms were developed for tracking thunderstorms *(Lakshmanan* and *Smith*, 2010). The TITAN technique (thunderstorm identification, tracking, analysis and nowcasting) was developed by *Dixon* and *Wiener* (1993), while *Morel et al.* (1997) used a procedure of the extent of overlaps. *Johnson et al.* (1998) applied projected centroid locations and *Han et al.* (2009) combined TITAN and overlap methods.

In Hungary, first investigations of severe convection from dynamical aspects were made in the 1960s (*Bodolainé et al.*, 1967; *Götz*, 1968). Later, remote sensing data appeared in the Hungarian studies (*Horváth* and *Práger*, 1985; *Boncz et al.*, 1987). Supercell thunderstorms and formation of tornadoes were described by *Horváth* (1997). Since the 2000s, nowcasting methods (*Horváth* and *Geresdi,* 2003; *Horváth et al.,* 2007; *Csirmaz et al.*, 2013) and radar-based thunderstorm climatology (*Horváth et al.,* 2008) have appeared as well.

The aim of this study is to survey incidence and behavior of severe thunderstorms in Hungary using radar observations in the period of 2004–2012 applying radar-reflectivity based recognition and tracking methods.

2. Methodology

Hungary is covered by three weather radars operated by the Hungarian Meteorological Service. These locators work in the western part (near the western end of Lake Balaton on the hill Pogányvár), in the central region (in the southeastern district of Budapest) and in the eastern part (in Nyíregyháza-Napkor) of the country; this system covers most part of the Carpathian Basin. All of them are C-band Doppler radars (wave length=5 cm) (*Geresdi*, 2004) and have been working in operative mode since 2004. During the operative measurement, the Doppler-wind was applied for noise filtering and the results were filtered and smoothed into composite fields. From each scan the highest reflectivity data of the relevant vertical column were chosen and placed into the composite image (*Collier,* 1996). The resolution of the composite PPI (plan position indicator) images was originally 2×2 km in space and 15 minutes in time.

The first step was making a more accurate cell tracking method: the original radar image frequency had to be increased from 15 minutes to 1 minute. The TREC method (tracking radar echoes by correlation) *(Tuttle* and *Foote,* 1990; *Horváth et al.,* 2012) was applied to do the time interpolation. During TREC procedure, correlations were searching between two consecutive radar images, and motion vectors were calculated which describe displacement of radar echoes. The computed motion vector field was used for calculating series of "artificial" radar images with 1 minute frequency. A brief description of TREC is given in *Appendix A.* For further noise reducing of reflectivity, median-filter method (*Tukey*, 1977) was also applied before beginning the analysis.

The second step was the thunderstorm identification using calculated 1 minute frequency radar reflectivity images. TITAN method developed by *Dixon* and *Wiener* (1993) represents irregular shaped thunderstorms by regular, best fitting ellipses. Parameters of an ellipse can be objectively used to describe the place (by coordinates of ellipses center), the size (by area of ellipses), and even the shape (minor and major axes) of a thunderstorm. The mathematical background of TITAN is given in *Appendix B.* Ellipses calculation of TITAN method expects two types of thresholds as input data: reflectivity limit (R_{min}) and area limit (N_{min}). R_{min} shows the minimum reflectivity of radar pixels that are involved into ellipse calculation. N_{min} shows the minimum number of pixels whose reflectivity values have to be equal or larger than R_{min}, and the strong echoes must form a continuous area. For example: R_{min}= 45 dBZ and N_{min}=5 mean that the procedure orders ellipses to such thunderstorm cells which have at

least 5 pixels with higher than 45 dBZ reflectivity. These calculated ellipses were named *thunderstorm ellipses* (*Horváth et al.*, 2008). An example of cell detecting is shown on *Fig. 1,* where R_{min} is set to 45 dBZ and N_{min}=5 pixels. The results can be visualized by the Hungarian Advanced Workstation (HAWK) system (*HMS*, 2012). Note, that the applied radar resolution is 2 km in space, thus N_{min}=5 means that the size of a thunderstorm ellipse has to be equal or greater than 20 km^2. In the present study, N_{min} was assumed to 5 radar pixels, and 3 different radar reflectivity thresholds were applied: 45, 50, and 55 dBZ. These objects were named *severe*, *highly severe*, and *extremely severe thunderstorm ellipses*, respectively. Using these high reflectivity values, the detected cells could be considered as models of severe thunderstorms and the small or weak convective cells were eliminated.

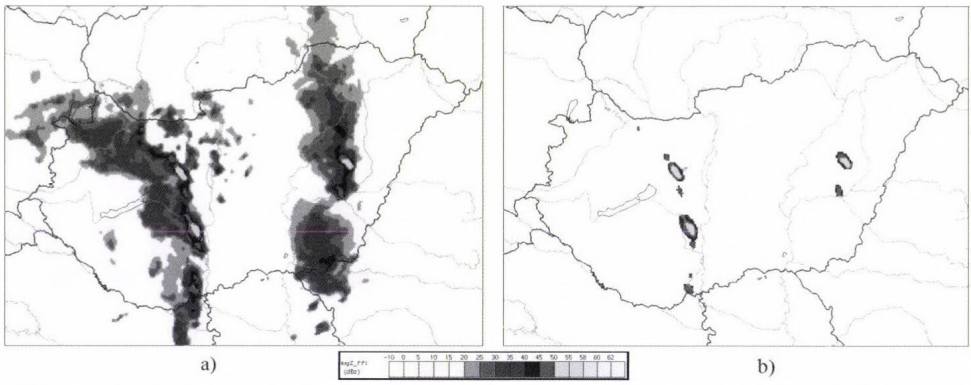

Fig. 1. Composite PPI radar images of thunderstorms observed on May 18, 2005, 16:00 UTC: a) original image b) image in which thunderstorms were represented by ellipses. For the visualization, the Hungarian Advanced Workstation (HAWK) system of the Hungarian Meteorological Service was used.

The third step was the cell tracking. For tracking thunderstorm ellipses, a special matching algorithm was applied where position of centrals and area of ellipses were considered to recognize the same cell in the next time step. The algorithm scanned the surrounding area of each ellipsis, and if one or more objects were found on the next image it chose the closest by distance and size. Also distance and size were used to recognize merging, splitting of thunderstorm ellipses. The 1 minute time steps made this procedure fairly reliable: verification on a 50-cell sample showed 98%, comparing with manually tracking. With these methods all information, i. e., size, lifetime, reflectivity, and the track of ellipses were followed and calculated from forming to dissipating and their tracks could be visualized.

The movement of severe thunderstorms, especially supercells, often shows deviation from the straight line, right or left turning of cells are important indicators of supercells (*Lemon* and *Doswell*, 1979). To describe these phenomena, the total deviation angle was introduced. This parameter was calculated in the following way: the displacement of a thunderstorm ellipse in the first 15 minutes designates a direction. Considering the direction coming from the next 15-minute interval, a deviation angle can be calculated. Summing up these angles during the lifetime of the thunderstorm ellipse, the total deviation angle was obtained *(Fig. 2.)*.

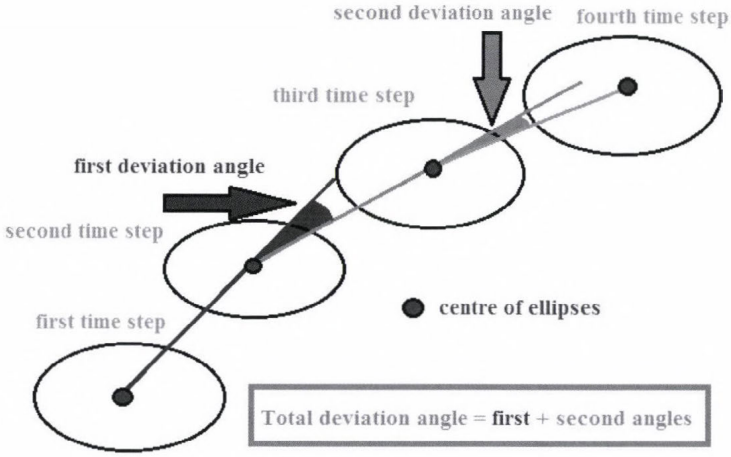

Fig. 2. Calculation method of the total deviation angle.

The present study focused on long-lived severe thunderstorms, so the investigation was applied only for severe, highly severe, and extremely severe thunderstorm ellipses with a lifetime more than 1 hour.

3. Results

In the 9-year period, 2625 severe, 597 highly severe, and 45 extremely severe long-lived ellipses were found. *Fig. 3* shows the distribution of lifetime for each type. The maximum of lifetime resolution was around 60 minutes which was followed by rather rapid than moderate decrease in number, and only few severe or highly severe cells lived more than 270 minutes. The average values were 100 minutes for severe, 95 minutes for highly severe, and 86 minutes for extremely severe ellipses.

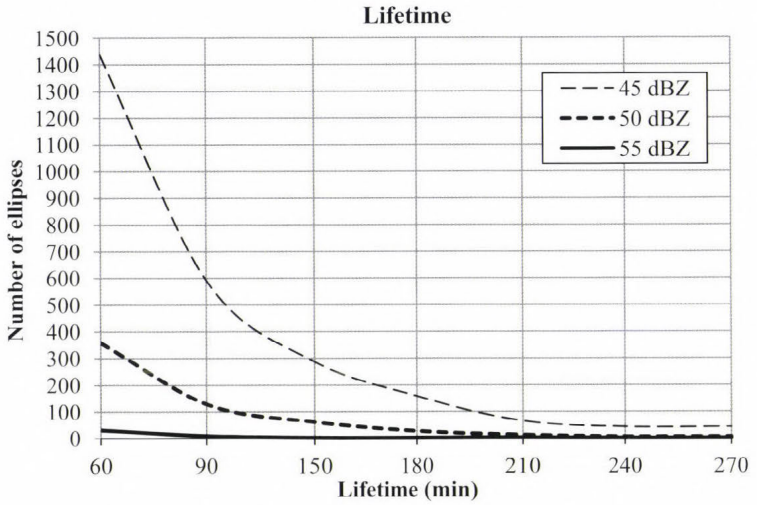

Fig. 3. Lifetime characteristics of long-lived severe (45 dBZ), highly severe (50 dBZ), and extremely severe (55 dBZ) ellipses.

The lengths of storm tracks had asymmetric distribution drawing an analogy to log-normal (*Fig. 4*). Most of investigated thunderstorms moved distance between 40 and 80 km during their lifetime and the average distance was about 70 km. There were some cells with extreme long lifetime. For example, on July 14, 2008 a strong cell crossed the whole domain (about 600 km) moving from southwest to northeast.

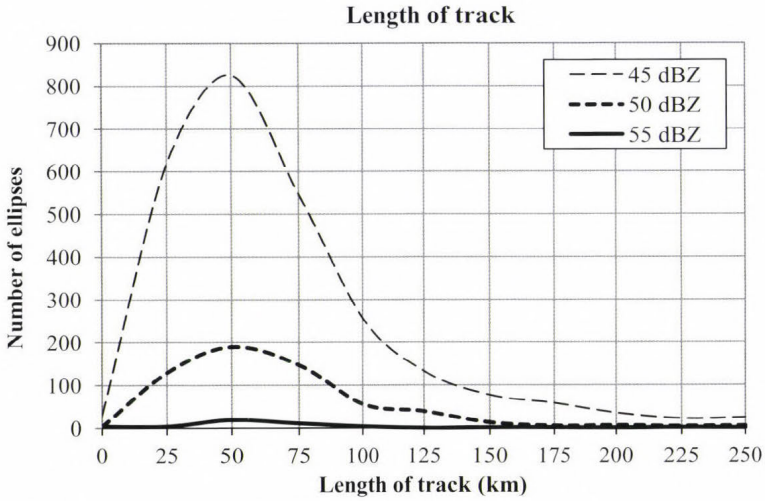

Fig. 4. Distribution of storm tracks' length for long-lived severe (45 dBZ), highly severe (50 dBZ), and extremely severe (55 dBZ) ellipses.

The total deviation angle was also investigated. The distribution of total deviation angles was mostly symmetric, especially for ellipses with 45 or 50 dBZ. The ratios between right and left directions were about 50–50% for severe, and 51–49% for highly severe thunderstorm ellipses. At extremely severe thunderstorms this rate was 58–42%, so for most intensive objects, right-turning movement occurred somewhat more often (*Fig. 5*). Considering all investigated thunderstorms, 23% have deviation angle larger than 40° and 13% have larger than 60 degrees.

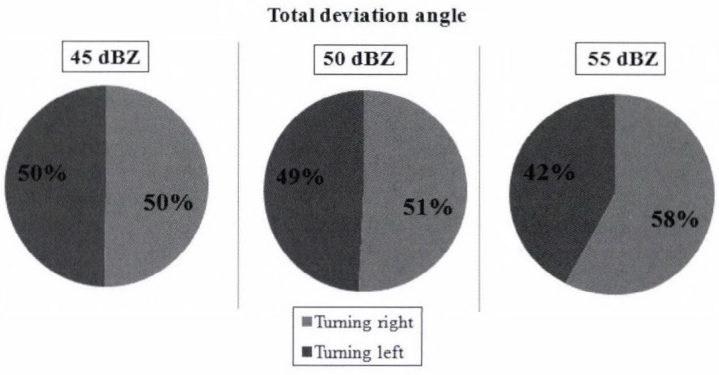

Fig. 5. Ratios between positive and negative deviation angles for severe (45 dBZ), highly severe (50 dBZ), and extremely severe (55 dBZ) ellipses.

Knowing lengths of tracks and lifetimes of thunderstorm-ellipses, their average speed could be calculated as well. The distributions of speed were asymmetric: the maximum was around 35–45 km/h (*Fig. 6*). Stronger cells moved faster. For severe thunderstorm ellipses the average speed was 42 km/h, for highly severe objects it was 45 km/h, and for extremely severe ellipses 49 km/h were calculated. Only few (8) severe objects' speed was higher than 100 km/h.

Numbers of severe thunderstorms had wide variability during 9 years. The highest values were in 2010 (with 611 severe, 146 highly severe, and 14 extremely severe ellipses). The lowest numbers were detected in 2005 (with only 95 severe, 13 highly severe, and 0 extremely severe objects).

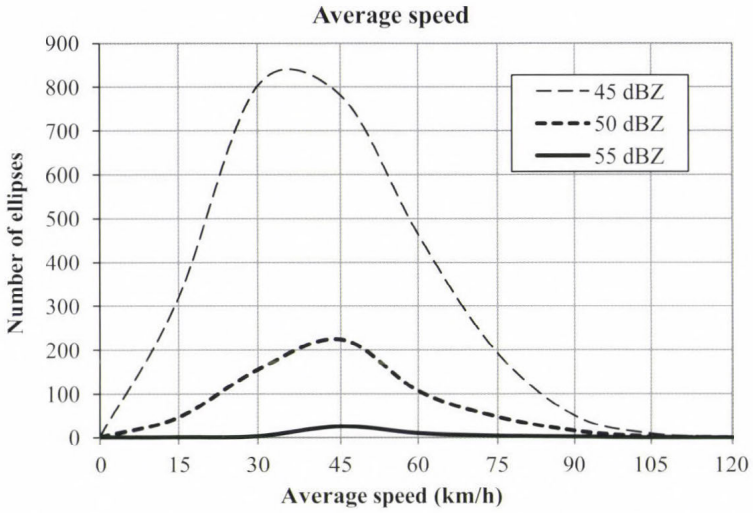

Fig. 6. Distribution of average speed for severe (45 dBZ), highly severe (50 dBZ), and extremely severe (55 dBZ) ellipses.

4. Conclusions

This study described a radar-based climatology of long-lived thunderstorms in Hungary focusing on cell motion and tracking in the period of 2004–2012. In our research, intensive thunderstorms were represented with ellipses calculated by the TITAN identification method. During the investigation, severe, highly severe and extremely severe thunderstorm ellipses were distinguished using thresholds of 45 dBZ, 50 dBZ, and 55 dBZ. The investigation was restricted only to thunderstorm ellipses with lifetime exceeded 1 hour.

In the 9-year period, 2625 severe, 597 highly severe, and 45 extremely severe long-lived ellipses were found. The time resolution of thunderstorms during this 9-year period was rather changeable showing the role of larger, synoptic-scale phenomena that determinate condition of the convection. The average values of lifetime were 100 minutes for severe, 95 minutes for highly severe, and 86 minutes for extremely severe ellipses. These values are promising from nowcasting point of view, because cell tracking is a relative easy technique, and persistent cells can be forecasted by extrapolation more than an hour ahead. Lengths of tracks were mostly between 40 and 80 km. Supposing that these storms are able to cause hail and wind damages, there is a high probability that populated places, sensitive infrastructure, or industrial areas were hit by the investigated storms. Stronger cells moved faster, the average speeds were 42 km/h for severe, 45 km/h for highly severe, and 49 km/h for

extremely severe objects. At most intensive cells, right-turning movements were more frequent. At 23% of thunderstorms, significant deviations from the straight line (>40°) were found. These deviations suggest existence of supercells. Considering the 23% of 2625 detected thunderstorms in 9 years, *the estimated number of supercells in a year is 67.*

These results are parts of the work dealing with severe storm climatology. Further researches are needed involving surface observations to obtain wind, pressure, and precipitation fields associated with thunderstorms, and more NWP case studies should be used for understanding the mechanism of severe convective storms in the Carpathian Basin.

Acknowledgement: This study is related to COST-0905 project.

References

Battan, L.J., 1952: Formation of precipitation in natural clouds indicated by radar. Preprints, Third Conference on Radar Meteorology, Montreal, PQ, Canada, Amer. Meteor. Soc., A9–A16.

Bodolainé, J.E., Bodolai, I. and *Böjti, B.*, 1967: Macrosynoptical conditions for the formation of Slovenenian squall lines and some properties of cold fronts with thunderstorm. *Időjárás 67*, 129–143.

Boncz, J., Kapovits, A., Pintér, F. and *Tänczer, T.*, 1987: A method for the complex analysis of synoptic weather radar and satellite data. *Időjárás 91*, 11–22.

Browning, K. A., 1964: Airflow and precipitation trajectories within severe local storms which travel to the right of the winds. *J. Atmos. Sci. 21*, 634–639.

Collier, C.G., 1996: Application of weather radar system: A Guide to uses of radar in meteorology and hydrology. John Wiley & Sons.

Csirmaz, K., Simon, A., Pistotnik, G., Polyánszky, Z., Neštiak, M., Nagykovácsi, Zs. and *Sokol, A.*, 2013: A study of rotation in thunderstorms in a weakly- or moderately-sheared environment. *Atmos. Res. 123*, 93–116.

Dixon, M., and *Wiener, G.*, 1993: TITAN: Thunderstorm Identification, Tracking, Analysis and Nowcasting—A radar-based methodology. *J. Atmos. Oceanic Technol. 10*, 785–797.

Doviak, R.J. and *Zrinic, D.S.*, 1993: Doppler Radar and Weather Observations. Academic Press, 335–340.

Geresdi, I., 2004: Felhőfizika. Dialóg Campus, Budapest, 153–170. (in Hungarian)

Götz, G., 1968: Hydrodynamic relationships between heavy convection and the jet stream. *Időjárás 72*, 157–165.

Han, L., S. *Fu,* L. *Zhao,* Y. *Zheng,* H. *Wang,* and *Lin, Y.*, 2009: 3D convective storm identification, tracking and forecasting–An enhanced TITAN algorithm. *J. Atmos. Oceanic Technol. 26*, 719–732.

Henry, S.G., 1993: Analysis of thunderstorm lifetime as a function of size and intensity. Preprints, 26th Conf. on Radar Meteorology, Norman, OK, Amer. Meteor. Soc., 138–140.

Horváth, Á.,1997: Tornado. *Légkör 62*, 2–9. (in Hungarian)

Horváth, Á. and *Práger, T.*, 1985: Zivatarláncok dinamikája és előrejelezhetőségét. *Időjárás 89*, 141–160. (in Hungarian)

Horváth, Á. and *Geresdi, I.*, 2003: Severe storms and nowcasting in the Carpathian Basin. *Atmos. Res. 67–68*, 319–332.

Horváth, Á., Ács, F. and *Seres, A. T.*, 2008: Thunderstorm climatology analyses in Hungary using radar observations. *Időjárás 112*, 1–13.

Horváth, Á., Seres, A.T. and *Németh, P.*, 2012: Convective systems and periods with large precipitation in Hungary. *Időjárás 116*, 77–91.

Horváth, Á., Geresdi, I., Németh, P. and *Dombai, F.*, 2007: The Constitution Day storm in Budapest: Case study of the August 20, 2006 severe storm. *Időjárás 111*, 41–65.

HMS, Hungarian Meteorological Service, 2012: HAWK-3 visualization system. http://www.met.hu/en/omsz/tevekenysegek/hawk/

Johnson, J., P. *MacKeen*, A. *Witt*, E. *Mitchell*, G. *Stumpf*, M. *Eilts*, and *Thomas, K.*, 1998: The Storm Cell Identification and Tracking algorithm: An enhanced WSR-88D algorithm. *Weather Forecast 13*, 263–276.

Klemp, J.B., 1987: Dynamics of tornadic thunderstorms. *Ann. Rev. Fluid Mech. 19*, 369–402.

Lakshmanan, V. and *Smith, T.*, 2010: An Objective Method of Evaluating and Devising Storm-Tracking Algorithms. *Weather Forecast. 25*, 701–709.

Lemon, L. R. and *Doswell III, C. A.*, 1979: Severe thunderstorm evolution and mesocyclone structure as related to tornadogenesis. *Mon. Weather Rev. 107*, 1184–1197.

MacKeen, P.L., Brooks, H.E. and *Elmore, K.L.*, 1999: Radar Reflectivity–Derived Thunderstorm Parameters Applied to Storm Longevity Forecasting. *Weather Forecast. 14*, 289–295

Morel, C., *Orain, F.*, and *Senesi, S.*, 1997: Automated detection and characterization of MCS using the Meteosat infrared channel. Proc. Meteorological Satellite Data Users Conf., Brussels, Belgium, EUMETSAT, 213–220.

Tukey, J.W., 1977: Exploratory Data Analysis. Addison-Wesley, Reading, 688. pp.

Tuttle, J.D. and *Foote, B.*, 1990: Determination of the Boundary Layer Airflow from Single Doppler Radar. *J.Atmos. Ocean. Tech. 7*, 218–232.

Wilhelmson, R.B. and *Klemp J.B.*, 1978: A numericalsStudy of storm splitting that leads to long-lived storms. *J. Atmos. Sci. 35*, 1974–1986.

Wilson, J.W., 1966: Movement and predictability of radar echoes. Tech. Memo IERTM-NSSL-28, National Severe Storms Laboratory.

Appendix A

Calculation of motion vectors using time series of radar reflectivity

(*Tuttle* and *Foote*, 1990; *Horváth et al.*, 2012)

In the period of 2004–2012, the Hungarian radar network collected data in 15-minute cycles. For a more accurate cell tracking procedure, the original radar image frequency had to be increased using the correlation tracking method TREC (tracking radar echoes by correlation; *Tuttle* and *Foote*, 1990). During the TREC procedure, the radar grid was divided into so-called macro grids, and the calculation of motion vectors was based on maximum correlations for the macro grids. After quality control to filter out noisy vectors on macro grids, fine resolution motion vectors were interpolated for all grid points of the original radar grid. Once a motion vector field – belonging to radar images at time T2 and T1 – is available, interpolation of the radar reflectivity can be done at any time between T1 and T2. Echoes from T1 are moved forward and echoes from T2 moved backward by motion vectors, and the reflectivity of a given pixel is interpolated between the forward and backward moving reflectivity values as shown in *Fig. 7.*

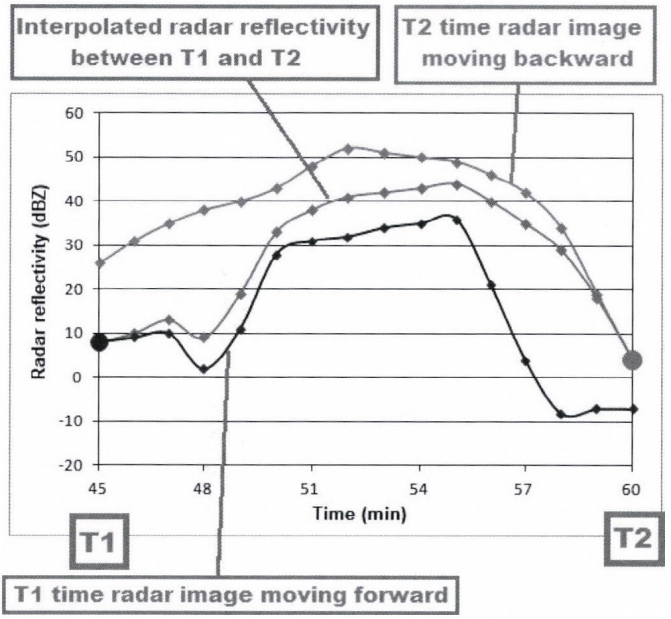

Fig. 7. Interpolation of radar reflectivity using motion vectors.

The application of the TREC method offers a more realistic and accurate cell tracking procedure. The optimal interpolation time step for calculation was found to be 1 minute.

Appendix B

Mathematical background of the identification

(*Dixon* and *Wiener*, 1993)

Suppose there is an irregular cluster on a radar image which has n detected pixels. The center of a cluster is defined by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i , \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i , \tag{1}$$

where x and y indicate the longitude and latitude of pixels which have reflectivity higher than a given threshold value. The covariance matrix of this cluster is

$$A = \mathrm{cov}_{xy} = \begin{bmatrix} d & e \\ e & f \end{bmatrix}, \tag{2}$$

where d is the deviation from the center by the x coordinate

$$d = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, \tag{3}$$

f is the deviation from the center by the y coordinate

$$f = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2, \tag{4}$$

and e is

$$e = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}). \tag{5}$$

The eigenvalues of the covariance matrix are given by

$$\lambda_1, \lambda_2 = \frac{(d+f) \pm \left[(d+f)^2 - 4(df - e^2)\right]^{1/2}}{2}. \tag{6}$$

The normalized eigenvectors of this matrix are

$$v = \left[\frac{1}{(1+g)^2}\right]^{1/2} \qquad \mu = -gv, \tag{7}$$

where

$$g = \frac{f + e - \lambda_1}{d + e - \lambda_2}. \tag{8}$$

Then the rotation of the ellipse major axis relative to the x axis is given by these vectors

$$\theta = \tan^{-1}\left(\frac{\nu}{\mu}\right). \tag{9}$$

The eigenvalues of the covariance matrix (λ_1 and λ_2) represent the variances of the data (pixels)

$$\sigma_{major} = \lambda_1^{1/2}, \qquad \sigma_{\min or} = \lambda_2^{1/2}. \tag{10}$$

The area of the detected cluster is

$$A = ndxdy \tag{11}$$

where dx and dy are the grid spacing on the radar image. The area of an ellipse is given by

$$T = \pi\, ab, \tag{12}$$

where a and b represents the major and minor axes of the ellipses.
The main idea is that the area of the irregular cluster and the area of the ellipse have to be equal, therefore

$$A = T. \tag{13}$$

So the major and minor axes of the ellipses can be calculated by

$$a = \sigma_{major}\left(\frac{A}{\pi\sigma_{\min or}\sigma_{major}}\right)^{1/2}, \quad b = \sigma_{\min or}\left(\frac{A}{\pi\sigma_{\min or}\sigma_{major}}\right)^{1/2}. \tag{14}$$

With these parameters ($\bar{x}, \bar{y}$, a, b, θ), the focus points and the equation of the ellipse can be determined.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 1, January – March, 2015, pp. 53–68

Estimating spectra of unevenly spaced climatological time series

István Matyasovszky

Department of Meteorology, Eötvös Loránd University,
Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary,
E-mail: matya@caesar.elte.hu

(Manuscript received in final form May 3, 2014)

Abstract—Spectral analysis is often based on a comparison of the periodogram and the spectral density of a so-called background noise. This spectral density is estimated by fitting a first order autoregressive (AR(1)) process to data, as climatological time series generally exhibit red noise spectra that can be approximated by AR(1) models. When periodogram exceeds some threshold at a frequency, the spectrum is said to differ from this background noise, and the frequency is characteristic for the time series in question. The traditional periodogram, however, must not be used without modifications for unevenly spaced data. Additionally, red noise, characterized by spectral densities monotone increasing to low frequencies, covers a much wider class of processes than the AR(1) processes. Our purpose is (1) to introduce a new periodogram (ELSP) based on a least square (LS) fit for an entire set of frequencies instead of using the well-known Lomb-Scargle periodogram (LSP) based on individual LS fits for individual frequencies; (2) to estimate the spectral density without any assumption on its analytical form using the nearly isotonic regression (NIR) method with either ELSP or LSP. As NIR allows the possibility of deviations from red noise, comparison of the periodogram with a background noise is unnecessary. Note that ELSP has never been used before as is a new concept for defining the periodogram for unevenly spaced data. NIR is more or less known for curve fitting problems but has not been applied yet to spectral density estimation. Three examples show that although ELSP does not radically differ from LSP, NIR-ELSP and NIR-LS spectra can exhibit distinct shapes.

Key-words: spectra, unevenly spaced data, Lomb-Scargle periodogram, red noise, nearly isotonic regression

1. Introduction

Literature of spectral analysis of climatological time series is extremely broad. The task of the spectral analysis is to identify sets of frequencies that essentially contribute to the behavior of time series. A common way is to calculate the

periodogram and then to fit a first order autoregressive (AR(1)) process to data in order to model the so-called background noise. When periodogram exceeds some threshold at a frequency, the spectrum is said to differ from this background noise, and the frequency is characteristic for the time series in question. The threshold depends on the AR(1) model and the significance level selected.

Let $x(t_1),...,x(t_n)$ be a stationary time series observed at $t_1,...,t_n$. Usually, the data set is evenly spaced, and $t_i = i$ can be taken. Hence, the time series can be written as $x_1,...,x_n$. The background noise is taken as red noise, and is generally described with the spectral density

$$f(\lambda) = (\sigma_e^2 / \pi)/(1 + a^2 - 2a\cos(\lambda)), \sigma_e^2 = \sigma^2(1 - a^2) \qquad (1)$$

of an AR(1) process with substituting the autoregressive parameter a and variance σ^2 with their consistent estimates $\hat{a}$ and $\hat{\sigma}^2$ obtained from $x_1,...,x_n$. As Eq. (1) provides red noise spectra under positive a, red noise and AR(1) spectra are seldom used as synonyms. But red noise, characterized by spectral densities monotone increasing to low frequencies, represents a much wider class of processes than the AR(1) processes, and the usage of AR(1) spectra can thus fail to properly detect frequencies mainly contributing to spectra. Therefore, we propose a method that ignores the comparison of periodograms with background noise models.

In some cases, time series are unevenly spaced, and hence the periodogram defined for evenly spaced data must not be used without modifications. There are two main ways to handle the problem. The first one is based on an interpolation of the data onto an equispaced time grid, and this new regularly spaced data set is analyzed with traditional techniques. Such data manipulations, however, always deform the true spectra (*Broersen*, 2006). The other way, which is addressed in this paper, produces a reformulated periodogram directly from data. Evidently, such a periodogram is always affected by temporal distributions of data spacing. Generally, the Lomb-Scargle periodogram (LSP) (*Lomb*, 1976; *Scargle*, 1982), based on a simple least squares (LS) estimation procedure, is used for the purpose. However, important statistical properties of LSP, e.g., its probability distribution are known only for white background noise. Additionally, the bias of LSP for unevenly spaced data can be substantially higher than that of the periodogram for evenly spaced data (*Vio et al.*, 2010), principally at high frequencies (*Schulz* and *Mudelsee*, 2002). Therefore, we will examine first the properties of LSP. Then we will propose a periodogram that is based on a so-called entire least squares (ELS) technique (*Matyasovszky,* 2013a). Properties of this newly introduced periodogram will also be discussed.

The methodology is described in Sections from 2 to 4. As paleoclimatological records represent a typical case of unevenly spacing, our technique is demonstrated by three paleoclimatic records in Section 5. Finally, a section for discussion and conclusions is provided.

2. Periodograms

2.1. Lomb-Scargle periodogram (LSP)

Let $x(t_1), x(t_2), ..., x(t_n)$ with $t_1 = 1, t_n = N$ be a time series coming from a stationary stochastic process with mean zero. The LSP for any frequency λ_j in the interval $(0, \pi]$ is based on an LS procedure as follows. Parameters a_j and b_j that minimize

$$\sum_{i=1}^{n} (x(t_i) - a_j \cos(\lambda_j t_i) - b_j \sin(\lambda_j t_i))^2$$

are obtained with the solution of the system of equations

$$\underline{\underline{D}}\,\underline{c} = \underline{w}, \tag{2}$$

where $\underline{\underline{D}} = \underline{\underline{Z}}^T \underline{\underline{Z}}$, and the elements of $\underline{\underline{Z}}$ are $z_{i1} = \cos(\lambda_j t_i)$, $z_{i2} = \sin(\lambda_j t_i)$, $i = 1, ..., n$, furthermore $\underline{w} = \underline{\underline{Z}}^T \underline{x}$, and $\underline{x} = (x(t_1), ..., x(t_n))^T$, $\underline{c} = (a_j, b_j)^T$ with superscript T denoting transpose. The quadratic form

$$I_{LS}(\lambda_j) = 1/(2\pi)\,\underline{c}^T \underline{\underline{D}}\,\underline{c} \tag{3}$$

defines LSP. For evenly spaced data, Eq. (3) becomes to the well-known expression

$$I(\lambda_j) = n/(4\pi)(a_j^2 + b_j^2) \tag{4}$$

with

$$a_j = 2/n \cdot \sum_{i=1}^{n} x_i \cos(\lambda_j i), \quad b_j = 2/n \cdot \sum_{i=1}^{n} x_i \sin(\lambda_j i).$$

It is known from the LS procedure that $\underline{c}$ is asymptotically normally distributed under very general conditions. Supposing that $\underline{x}$ comes from a white

noise process, the covariance matrix of $\underline{c}$ is $\sigma^2 \underline{\underline{D}}^{-1}$, where σ^2 is the variance of the process. Hence, the random variable $I_{LS}(\lambda_j)$ defined with the quadratic form in Eq. (3) follows an exponential distribution (*Scargle*, 1982). The probability distribution of LSP when $\underline{x}$ does not come from a white noise process is, however, an open question. Therefore, we will discuss this issue together with the expected value of $I_{LS}(\lambda_j)$ in Section 4. Although λ_j can be any of frequencies in an interval $[\lambda_{\min}, \lambda_{\max}]$, it is advisable to define a grid $2\pi j/(n\Delta), j = 1,..., L$, where L is the largest integer not larger than $n/2$, $\lambda_{\max} = \pi/\Delta$ is the average Nyquist frequency (*Stoica et al.*, 2009), and Δ is the average of time steps $\delta_i = t_i - t_{i-1}, i = 2,..., n$. As the frequency range of $I(\lambda)$ is $(0, \pi]$, the range $[\lambda_{\min}, \lambda_{\max}]$ of $I_{LS}(\lambda)$ is generally rescaled into $[2\pi/n, \pi]$ for convenience. $I_{LS}(\lambda)$ can thus be viewed as $I(\lambda)$ of a time series sampled evenly at time steps Δ. An important difference is, however, that elements of $I_{LS}(\lambda_j)$ are correlated in contrast to elements of $I(\lambda_j)$. Additionally, the bias of $I_{LS}(\lambda_j)$ can be higher than the bias of $I(\lambda_j)$ (*Vio et al.*, 2010). This is due to the interrelationship between unevenly spacing and the effect of the omission of frequencies different from λ_j when calculating $I_{LS}(\lambda)$ at λ_j. Furthermore, it is known that $I(\lambda)$ integrates to the sample variance $\hat{\sigma}^2$ in the sense that

$$\frac{2\pi}{n}\sum_{j=1}^{L} I(\lambda_j) = \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x(t_i) - \bar{x})^2,$$

while $I_{LS}(\lambda)$ does not integrate to $\hat{\sigma}^2$.

2.2. Entire least squares periodogram (ELSP)

The deficiency of LSP is that it handles the different frequencies separately. Therefore, we propose an entire least squares (ELS) procedure by calculating the constants a_j, b_j at frequencies $2\pi j/(n\Delta), j = 1,..., L$ simultaneously. This results in a system of equations Eq. (2), but with

$$z_{ij} = \left\{ \begin{matrix} \cos(\lambda_j t_i), j = 1,..., L \\ \sin(\lambda_j t_i), j = L+1,..., 2L \end{matrix} \right\}, \quad i = 1,..., n, \ \underline{c} = (a_1,..., a_L, b_1,..., b_L)^T.$$

It is easy to see that $\underline{c}^T \underline{\underline{D}}\underline{c} = n\hat{\sigma}^2$. As $\underline{\underline{D}}\underline{c} = \underline{w}$, hence $\underline{c}^T \underline{w} = c_1 w_1 + ... + c_{2L} w_{2L} = n\hat{\sigma}^2$. Therefore, we define ELS periodogram (ELSP) as the contribution of frequencies to the sample variance, i.e.,

$$I_{ELS}(\lambda_j) = 1/(2\pi)(c_j w_j + c_{j+L} w_{j+L}). \quad (5)$$

Although every frequency λ_j is affected by frequencies not involved in the estimation procedure, the accuracy of ELSP is expected to be higher than the accuracy of LSP. This is because ELSP is defined for the entire set of frequencies $\lambda_j, j = 1,...,L$ and not for particular frequencies separately. Properties of ELSP will be discussed in Section 4.

3. Estimating spectra

A stationary stochastic process exhibits red noise spectrum when its spectral density function satisfies $f(\lambda) \geq f(\omega), \lambda < \omega$ for every $\lambda, \omega \in (0, \pi]$. Estimation of a spectral density corresponding to this definition is now based on an LS technique. The solution of the LS problem

$$\min\left\{\sum_{j=1}^{L}\left(I_*(\lambda_j) - \hat{f}(\lambda_j)\right)^2\right\}, \quad \hat{f}(\lambda_j) \geq \hat{f}(\lambda_k),\ j < k$$

can be obtained using a procedure called isotonic regression (IR). Namely,

$$\hat{f}(\lambda_j) = \min_{i \leq j} \max_{j \leq k} \frac{I_*(\lambda_i) + ... + I_*(\lambda_k)}{k - i + 1}$$

for $j = 1,...,L$, and $\hat{f}(\lambda)$ is left-continuous otherwise (*Zhao* and *Woodroofe*, 2012), where I_* can be either I, I_{LS}, or I_{ELS}. Note that $\hat{f}(\lambda)$ is monotone decreasing and is stepwise constant over certain frequency ranges.

Behavior of periodogram elements at frequencies close to possibly existing discrete frequencies (frequencies contributing to discrete spectra) substantially differs from the behavior of the majority of periodogram elements. Thus, periodogram elements at these frequencies should be taken as outliers, and an IR, robust against outliers has to be found. *Álvarez* and *Yohai* (2011) proposed a robust IR technique that can thus be used as a method to estimate the spectral density of the red background noise without any assumption on its analytical form. A possibility to find essential frequencies is to detect significant deviations of the periodogram from this background noise utilizing the statistical properties of robust IR (*Matyasovszky*, 2013b). Another way is to give up the background noise concept and estimate the spectral density non-robustly and without monotonicity constraint. The nearly-isotonic regression (NIR) introduced by *Tibshirani et al.* (2011) permits the possibility of deviations from

monotonicity when necessary. The necessity of monotonicity violations is controlled via a parameter β that is estimated within the procedure. The task is to find $\hat{f}(\lambda_1),...,\hat{f}(\lambda_L)$ minimizing the quantity

$$\left\{\sum_{j=1}^{L}\left(I_*(\lambda_j)-\hat{f}(\lambda_j)\right)^2+\beta\sum_{j=1}^{L-1}\left(\hat{f}(\lambda_j)-\hat{f}(\lambda_{j+1})\right)_+\right\},$$

where $(u)_+ = -u$ when u is negative, and zero otherwise. When $\beta = 0$, the solution is $\hat{f}(\lambda_j) = I_*(\lambda_j)$, and letting $\beta \to \infty$, we obtain the isotonic regression. The optimal value of β, which can be estimated (*Tibshirani et al.,* 2011), gives a trade off between monotonicity and goodness-of-fit. The resulting $\hat{f}(\lambda)$ can thus correspond to red noise or colored noise with certain local peaks according to the shape of $f(\lambda)$.

4. Properties of periodograms

4.1. Lomb-Scargle periodogram (LSP)

Now we examine the probability distribution and the expected value of LSP. For this reason, take first a two-dimensional normal random vector $\underline{\xi}$ with expected value zero and covariance matrix $\underline{\underline{P}}$. Let $\underline{\underline{Q}}$ be a positive definite matrix of size (2x2), and take the random variable $q = \underline{\xi}^T \underline{\underline{Q}} \underline{\xi}$. This quadratic form can be rewritten as a linear combination of two independent chi-squared random variables with one degree of freedom, and the coefficients in this linear combination are the eigenvalues of $\underline{\underline{Q}}\underline{\underline{P}}$. If these coefficients are identical, q is distributed exponentially, but when difference between the coefficients is not too large, the probability distribution is also closely exponential (*Yuan* and *Bentler,* 2010).

It is known from the LS theory that covariance matrix of $\underline{c}$ is $\underline{\underline{D}}^{-1}\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}\underline{\underline{D}}^{-1}$ (*Nielsen*, 2011), where *(i,j)*th element of $\underline{\underline{B}}$ is the covariance between $x(t_i)$ and $x(t_j)$. As $\underline{c}$ is distributed asymptotically normally, we recognize that the quadratic form Eq. (3) asymptotically corresponds to the above mentioned random variable q with $\underline{\underline{P}} = \underline{\underline{D}}^{-1}\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}\underline{\underline{D}}^{-1}$ and $\underline{\underline{Q}} = 1/(2\pi)\underline{\underline{D}}$. When $\underline{x}$ comes from a white noise process, the eigenvalues of $\underline{\underline{Q}}\underline{\underline{P}} = 1/(2\pi)\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}\underline{\underline{D}}^{-1}$ are identical, and this is why LSP follows an exponential distribution under a white noise process. For other processes, the eigenvalues depend on length of the data set, spacing, autocovariances of the underlying process, and frequency. However, except for some degenerate cases, the difference between these eigenvalues is not too

large. Therefore, our final conclusion is that probability distribution of LSP is closely exponential even for processes different from white noise.

It can be shown (*Mathai* and *Provost*, 1992) that the expected value of q is the trace of $\underline{\underline{Q}}\underline{\underline{P}}$, and therefore, the expected value of Eq. (3) is the trace of

$$1/(2\pi)\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}\underline{\underline{D}}^{-1}.$$

Remember that trace of a quadratic matrix is the sum of diagonal elements of this matrix, which is identical with the sum of its eigenvalues. Utilizing basic trigonometric identities we obtain

$$E\left[I_{LS}(\lambda_j)\right]=trace\left(1/(2\pi)\sum_{k=-(N-1)}^{N-1}\underline{\underline{D}}_k\underline{\underline{D}}^{-1}B(k)\cos(\lambda_j k)\right), \tag{6}$$

where $B(k)$ is the autocovariance function of the process that generates the time series, $\underline{\underline{D}}_k=\underline{\underline{S}}_k^T\underline{\underline{S}}_k$ with *(i,j)*th element of $\underline{\underline{S}}_k$ $s_{k,ij}=\alpha_i\alpha_{i+|k|}(\cos(\lambda_j i),\sin(\lambda_j i))^T$, $i=1,...,N-|k|$, and $\alpha_i, i=1,..,N$ is an indicator series that equals with zero when no data is available at time *i*, while equals with one when data is available at time *i*. Evidently, Eq. (6) for evenly sampled data has the well-known form

$$E\left[I(\lambda_j)\right]=B(0)/\pi+2/\pi\sum_{k=1}^{n-1}(1-k/n)B(k)\cos(\lambda_j k), \text{ or} \tag{7}$$

$$E\left[I(\lambda_j)\right]=\int_{-\pi}^{\pi}g(\omega)K_n(\omega-\lambda_j)d\omega$$

with $g(\omega)=f(\omega),\omega\geq 0$, $g(\omega)=f(-\omega),\omega<0$, and

$$K_n(\omega)=1/(2\pi)\sum_{k=-(n-1)}^{n-1}u(k)\cos(\omega k),$$

where $u(k)=1-|k|/n$. Hence, $K_n(\omega)$ is the Fejér-kernel. In the frequency domain representation of

$$E\left[I_{LS}(\lambda_j)\right]=\int_{-\pi}^{\pi}g(\omega)K_n(\omega-\lambda_j)d\omega,$$

$u(k)$ corresponding to Eq. (6) is $u(k)=trace(\underline{\underline{D}}_k\underline{\underline{D}}^{-1})/2$. Note that LSP is an unbiased estimator for white noise processes (see Eq. (7)).

4.2. Entire least squares periodogram (ELSP)

As $\underline{\underline{D}}^{-1}\underline{w}=\underline{c}$, therefore $c_i w_i=(\underline{\underline{D}}^{-1}\underline{w})_i w_i$. Additionally, the covariance matrix of $\underline{w}$ is $\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}$. Utilizing these facts and basic trigonometric identities we obtain after *Mathai* and *Provost* (1992) that

$$E\left[I_{ELS}(\lambda_j)\right]=trace\left(1/(2\pi)\sum_{k=-(N-1)}^{N-1}\left\{\underline{\underline{D}}^{-1}\underline{\underline{D}}_k\right\}_j B(k)\cos(\lambda_j k)\right), \tag{8}$$

where $\left\{\underline{\underline{U}}\right\}_j$ denotes a matrix of size (2×2) consisting of *(j,j)*th, *(j,j+L)*th, *(j+L,j+L-1)*th *(j+L,j+L)*th elements of $\underline{\underline{U}}$. Here,

$$s_{k,ij}=\left\{\begin{matrix}\alpha_i\alpha_{i+|k|}\cos(\lambda_j i), j=1,...,L\\ \alpha_i\alpha_{i+|k|}\sin(\lambda_j i), j=L+1,...,2L\end{matrix}\right\}, i=1,...N-|k|.$$

Following *Yuan* and *Bentler* (2010) we observe that ELSP is also a linear combination of two independent chi-squared random variables with one degree of freedom, and if difference between coefficients in this linear combination is not very large, the probability distribution of ELSP is closely exponential. Eq. (8) for evenly sampled data also has the well-known form Eq. (7). In the frequency domain representation of

$$E\left[I_{ELS}(\lambda_j)\right]=\int_{-\pi}^{\pi}g(\omega)K_n(\omega-\lambda_j)d\omega,$$

$u(k)$ corresponding to Eq. (8) is $u(k)=trace\left(\left\{\underline{\underline{D}}^{-1}\underline{\underline{D}}_k\right\}_j\right)/2$. Note that for both the LSP and ELSP, $K_n(\omega)$ depends not only on n but also on data spacing and λ_j. However, ELSP is an unbiased estimator for white noise processes (see Eq. (8)).

5. Examples

5.1. Hallet Lake

The Hallet Lake temperature proxy record (*Mc Kay et al.*, 2008) from south-central Alaska available for a period of AD 492-2005 is based on biogenic silica preserved in the lacustrina sediments. Its data spacing varies from 1 to 35 years with an average spacing of 10.15 years. Data (n=150) are standardized to have zero mean and unit variance.

Fig. 1 shows no substantial differences between LSP and ELSP. However, many small differences count for much, as NIR performed with LSP (NIR-LSP) provides a red noise spectral density, while NIR with ELSP (NIR-ELSP) delivers a colored noise density. In this latter case, a moderate but wide peak of the spectral density reinforces the importance of multi-decadal oscillations shown in other Alaskan proxy records (e.g., *Wilson et al.*, 2007).

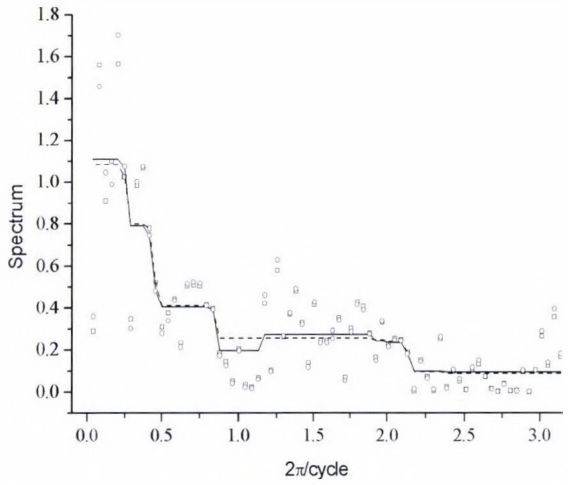

Fig. 1. ELSP (circle), LSP (square), NIR-ELSP spectral density (solid line), and NIR-LSP spectral density (dashed line) of Hallet Lake temperature proxy record AD 492-2005.

5.2. Vostok

Vostok ice core deuterium content data are available for the last 422,766 years (*Petit et al.*, 1999). Data spacing varies from 20 to 664 years with a mean spacing of 127.8 years. Data (n=3303) are standardized to have zero mean and unit variance. Not surprisingly, the highest peak of both the NIR-LSP and NIR-ELSP spectra appears at 105,500 years corresponding to the Earth eccentricity cycle (*Fig. 2*). The cycle related to obliquity can be seen at somewhat lower frequency (closer to the 41,000-year astronomical cycle) for NIR-ELSP than for NIR-LSP. More importantly, dominance of the eccentricity cycle is much clearer from NIR-ELSP, since NIR-ELSP peak at this cycle is substantially higher than the NIR-LSP peak. The ratio of the magnitude of peak at eccentricity to peak at obliquity is 1.63 for NIR-LSP, while it is 4.88 for NIR-ELSP. Peaks in an interval of 21,000-28,000 years corresponding to the precession are considerably more modest for NIR-ELSP than for NIR-LSP. The third largest peak of NIR-LSP is around a 60,000-year cycle which cannot be explained by a direct astronomical forcing but is probably a

side-effect of the aforementioned cycles (e.g., *Rial* and *Anaclerio*, 2000). Note, however, that this period is essentially missing in NIR-ELSP. *Fig. 3* shows peaks at high frequencies, too. Cycles around 270 years are substantially stronger with NIR-ELSP than with NIR-LSP. Similar fact can be mentioned for cycles corresponding to 500-525 years (almost double of the 270-year period). These cycles are clearly related to solar cycles listed in *Schove* (1983).

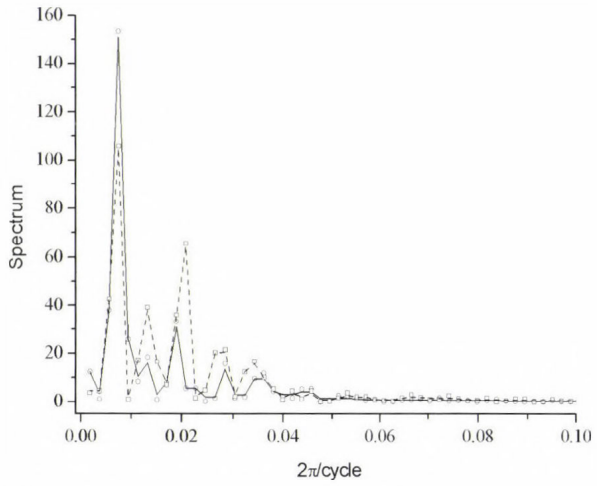

Fig. 2. ELSP (circle), LSP (square), NIR-ELSP spectral density (solid line), and NIR-LSP spectral density (dashed line) for Vostok ice core deuterium content data for the last 422,766 years at low frequencies.

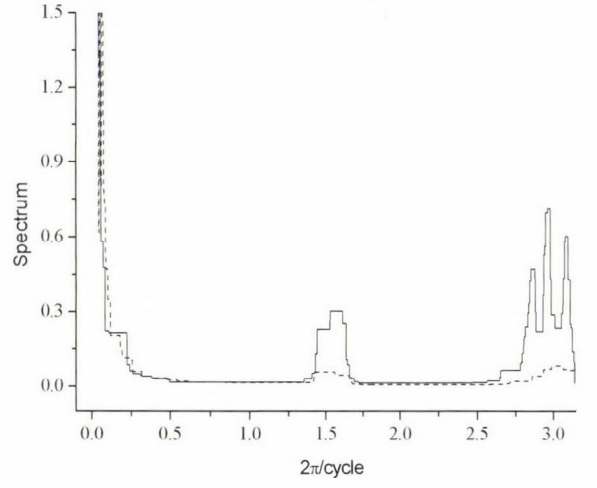

Fig. 3. NIR-ELSP spectral density (solid line) and NIR-LSP spectral density (dashed line) for Vostok ice core deuterium content data for the last 422,766 years at high frequencies.

5.3. GISP2

Oxygen-isotope data from GISP2 ice core from Greenland (*Groots* and *Stuvier*, 1997) are unevenly spaced in time, varying from 68 to 257 years with a mean spacing of Δ=125.8 years in the period between 15,000 and 60,000 BP. Data (n=358) are standardized to have zero mean and unit variance. *Schulz* and *Mudelsee* (2002) analyzed this record in order to detect a spectral peak at a 1470-year cycle corresponding to the spacing of the well-known Dansgaard-Oeschger events. Although the difference between ELSP and LSP does not seem substantial, NIR-LSP does not detect any spectral peak but does provide a red noise spectral density. In contrast, a peak at the above mentioned 1470-year cycle clearly appears when using the NIR method with ELSP (*Fig. 4*).

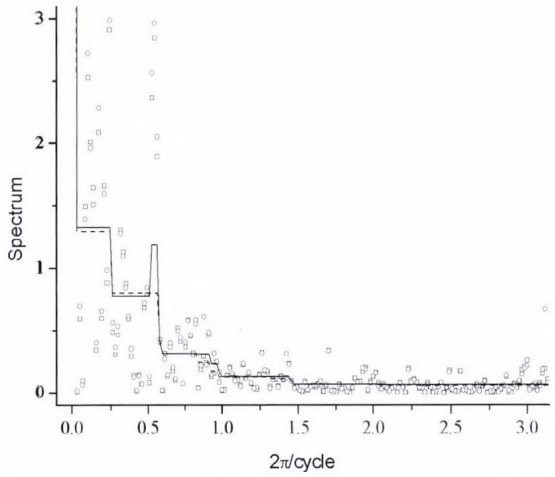

Fig. 4. ELSP (circle), LSP (square), NIR-ELSP spectral density (solid line), and NIR-LSP spectral density (dashed line) for GISP2 oxygen-isotope data for period between 15,000 and 60,000 BP. A big spectral peak around 40,000 years is not visible due to the vertical axis scale.

6. Discussion and conclusions

In order to demonstrate the ability of NIR and the drawback of AR(1) fitting, a simple example is taken by a stochastic process

$$Y_t = 0.5\cos(0.05\pi t) + X_t, \tag{9}$$

where the red background noise comes from a first order moving average (MA(1)) process $X_t = e_t - 0.5e_{t-1}$, and e_t is a white noise Gaussian process with $\sigma_e = 1$. Note that the variance corresponding to the discrete cycle $0.5\cos(0.05\pi t)$ is only 10% of the background noise variance. A time series of Y_t with n=400 is simulated and the spectrum is estimated with both the AR(1) fitting and the NIR method. The procedure is repeated 1,000 times. *Fig. 5* shows that the mean of the 1,000 NIR spectra exhibits a very sharp peak recognizing the discrete frequency at 0.05π. At the rest of frequencies, NIR spectrum reproduces well the background noise spectrum as compared to the background noise obtained with AR(1) fitting. Note that AR(1) spectral density is around two times higher than the true background spectral density at low frequencies causing difficulties in detecting the discrete frequency with traditional techniques. When omitting the discrete frequency from Eq. (9), the AR(1) spectral density is almost the same as in the previous case involving discrete frequency, while the NIR spectral density essentially coincides with the true background noise spectral density.

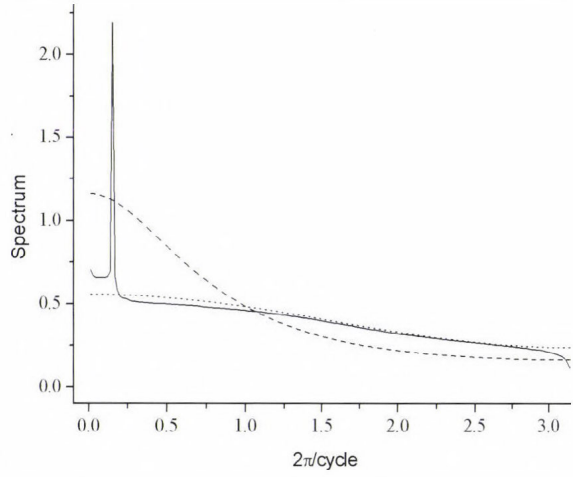

Fig. 5. Mean of NIR-ELSP spectral densities (solid line) and AR(1) spectral densities (dashed line) obtained from 1,000 simulated time series of length n=400 according to Eq. (9), and AR(1) spectral density (dotted line) of the background noise in Eq. (9).

Comparison of properties of LSP and ELSP is illustrated with the third data set in Section 5.3. As it was mentioned earlier, both the LSP and ELSP can be written as

$$I_*(\lambda_j) = h_1\xi_1 + h_2\xi_2, \tag{10}$$

where $I_*(\lambda_j)$ is either LSP or ELSP, and ξ_1 and ξ_2 are independent chi-squared random variables with one degree of freedom. For simplicity, dependence on the frequency of h_1 and h_2 is not indicated. If coefficients h_1, h_2 are identical, $I_*(\lambda_j)$ is distributed exponentially, but when the ratio $r = \max\{h_1, h_2\}/\min\{h_1, h_2\}$ is not too large, the probability distribution is closely exponential. Approximating the true probability distribution with an exponential distribution is highly accurate for ratios from r=1 to at least r=2-3 (*Yuan* and *Bentler*, 2010). Calculation of r (*Yuan* and *Bentler*, 2010) requires the autocorrelations of the underlying process. These are here substituted by autocorrelations corresponding to the AR(1) model fitted to data with a procedure described in *Schulz* and *Mudelsee* (2002). *Fig. 6* shows this ratio against frequencies. It is obvious that both periodograms can be taken as they are distributed exponentially. At very high frequencies, the distribution tends to deviate from the exponential one, but with smaller degree for ELSP than for LSP. Note that exponential approximation of the distribution of ELSP is accurate even at highest frequencies.

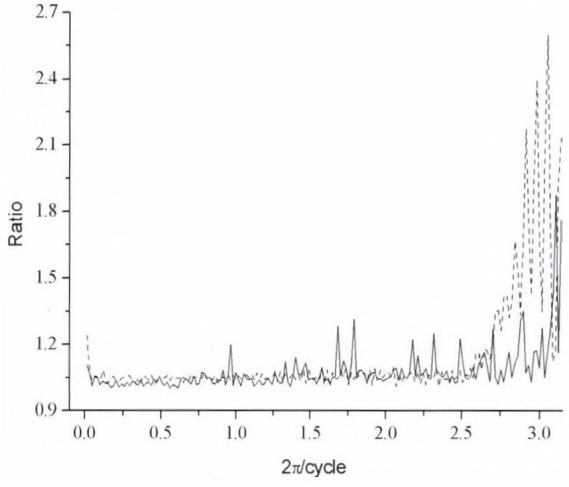

Fig. 6. Ratio $\max\{h_1, h_2\}/\min\{h_1, h_2\}$ against frequencies for ELSP (solid line) and LSP (dotted line), where coefficients h_1, h_2 are defined in Eq. (10).

Another important property of both the LSP and ELSP is that these periodogram elements at different frequencies are correlated for unevenly spaced data. *Fig. 7* shows the correlation between $I_{LS}(\lambda_j)$ and $I_{LS}(\lambda_i - \lambda_j)$ against the frequency λ_j and frequency shift $\lambda_i - \lambda_j$ for $i, j = 1,...,L$. It is apparent

that correlations are essentially zero at any frequencies and any frequency shifts except for the highest frequencies. At highest frequencies, the correlations are not negligible but only within narrow frequency shift intervals. The overall picture for ELSP (the corresponding figure is not shown) is the same but with slightly lower correlations. For instance, the largest correlation under every combination of λ_j and $\lambda_i - \lambda_j$ is 0.47 for LSP, while it is 0.43 for ELSP. These results are consistent with findings obtained for exponential approximation to the distribution of ELSP and LSP. The distribution of ELSP and LSP tends to deviate from the exponential distribution, when correlation between sinusoid and cosinusoid parts of the periodogram at a given frequency increases. Somewhat similar phenomenon can be found in *Vio et al.* (2010) but only for LSP and for time series simulated from white noise processes. It is to be mentioned that calculating the correlation between two periodogram elements utilizes that $\underline{c}$ has an asymptotic multivariate normal distribution. Hence, the mentioned correlation consists of fourth-order central moments of $\underline{c}$. These moments, due to the normality of $\underline{c}$, can be expressed via second order central moments, ie., via the covariance matrix of $\underline{c}$. Finally, this covariance matrix can be approximated using the autocorrelations corresponding to the AR(1) model fitted to data.

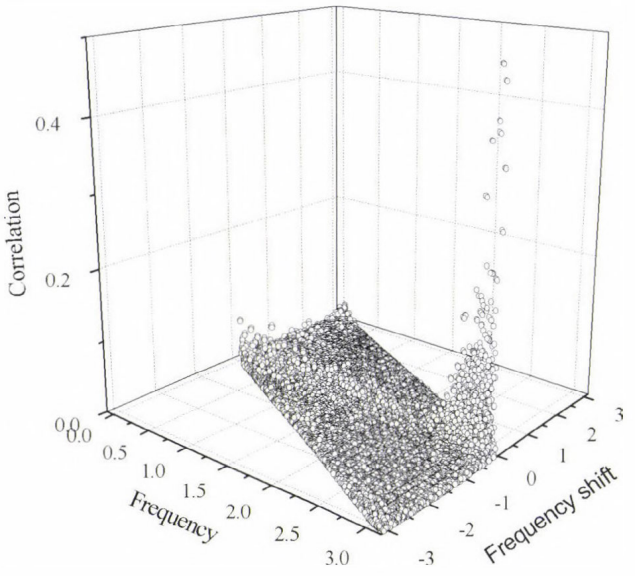

Fig. 7. Correlation between LSP elements at different frequencies and different frequency shifts.

Similar calculations shows that main conclusions discussed in the last two paragraphs are applicable for data sets of Sections 5.1 and 5.2.

Originally, LSP has been developed for time series generated by stochastic processes consisting of a certain number of periodic components plus a white noise process. Later, it has been using to estimate not only discrete spectra but spectral densities, too (e.g., *Schulz* and *Mudelsee*, 2002). Recognizing that LSP at a given frequency can be highly affected by other frequencies, *Stoica et al.* (2009) introduced a weighted least square fit at every separate frequency, where the weights are related to other LSP elements. The procedure thus necessitates an iterative technique requiring bigger computational effort than ELSP. More importantly, the weights are chosen with an approximation that holds accurately only for evenly spaced data. Such a simplification appears also in *Nygrén* and *Ulich* (2010). Here, after performing the entire least squares technique, the matrix $\underline{\underline{D}}$ is taken diagonal with elements $n/2$. Hence, their periodogram $n/(4\pi)(c_j^2 + c_{j+L}^2)$ provides a biased estimator even for white noise processes and does not integrate to $\hat{\sigma}^2$ (except for evenly spacing data) in contrast to our Eq. (5).

The matrix $\underline{\underline{D}}$ is generally close to being singular for large values of n, and solving Eq. (2) for $\underline{c}$ to form ELSP is not easy. Our experience is that traditional techniques such as Gauss-Seidel, successive over-relaxation, or conjugate gradient methods might be unsuccessful. Therefore, Eq. (2) was solved with a Monte Carlo technique (e.g., *Liu*, 2001), as this procedure is carried out numerically with a totally different scheme than the previous techniques.

References

Álvarez, E.E., and *Yohai, V.J.,* 2011: M-estimators for Isotonic Regression. arXiv: 1105.5065v1stat.ME.

Broersen, P.M.T., 2006: Autoregressive Order Selection for Irregularly Sampled Data. *IEEE Trans. Instrum. Meas. 54*, 1004–1009.

Groots, P.M. and *Stuvier, M.*, 1997: Oxygen 18/16 variability in Greenland snow and ice with 10^{-3}–10^{5}-year time resolution. J. *Geophys. Res. 102 (C12)*, 26455–26470.

Liu, J.S., 2001: Monte Carlo strategies in Scientific Computing. Springer, New York.

Lomb, N.R., 1976: Least-squares frequency analysis of unequally spaced data. *Astrophys. Space Sci. 39*, 447–462.

Mathai, A.M. and *Provost, S.B.,* 1992: Quadratic Forms in Random Variables. Taylor & Francis.

Matyasovszky, I., 2013a: Spectral analysis of unevenly spaced climatological time series. *Theor. Appl. Climatol 111*, 371–378.

Matyasovszky, I., 2013b: Estimating red noise spectra of climatological time series. *Időjárás 117*, 187–200.

Mc Kay, N.P., Kaufman, D.S., and *Michelutti, N.*, 2008: Biogenic-silica concentration as a high-resolution, quantitative temperature proxy at Hallet Lake, south-central Alaska. *Geophys. Res. Lett. 35*, LO5709.

Nielsen, A.A. ,2011: Least Squares Adjusments: Linear and Nonlinear Weighted Regression Analysis. Lecture Note, Lyngby, Denmark: IMM, DTU, http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=2804.

Nygrén, T., and *Ulich, Th.*, 2010: Calculation of signal spectrum by means of stochastic inversion. *Ann Geophys 28*, 1409–1418.

Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J.M., Basile, I., Bender, M., Chappellaz, J., Davis, J., Delaygue, G., Delmotte, M., Kotlyakov, V.M., Legrand, M., Lipenkov, V., Lorius, C., Pépin, L., Ritz, C., Saltzman, E., and *Stievenard M.*, 1999: Climate and Atmospheric History of the Past 420,000 years from the Vostok Ice Core, Antarctica. Nature 399, 429-436.

Rial, J.A., and *Anaclerio, C.A.*, 2000: Understanding nonlinear responses of the climate system to orbital forcing. *Quat. Sci. Rev. 19*: 1709–1722.

Scargle, J.D., 1982: Studies in astronomical time series analysis II: statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J. 261*, 835–853.

Schove, D.J. (Ed.), 1983: Sunspot cycles. *Benchmark Papers in Geology 68*, Hutchinson.Ross Publ.Co.

Schulz, M., and *Mudelsee, M.*, 2002: REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatological time series. *Comput. Geosci. 28*, 421–426.

Stoica, P., Li, J., and *He, H.*, 2009: Spectral Analysis of Nonuniformly Sampled Data: A New approach Versus the Periodogram. *IEEE Trans. Signal Process. 57*, 1415–1425.

Tibshirani, R.J., Hoefling, H., and *Tibshirani, R.*, 2011: Nearly-Isotonic Regression. *Technometrics 53*, 54–61.

Vio, R., Adreani, P., and *Biggs, A.*, 2010: Unevely-sampled signals: a general formalism of the Lomb-Scargle periodogram. *Astron. Astrophys. 519*, A86, 12.

Wilson. R., Wiles, G., D'Arrigo, R., and *Zweck, C.*, 2007: Cycles and shifts: 1,300 years of multi-decadal temperature variability in the Gulf of Alaska. *Clim. Dyn. 28*, 425–440.

Yuan, K-H., and *Bentler, P.M.*, 2010: Two simple approximations to the distributions of quadratic forms. *Br J Math Stat Psychol* 63, 273–291.

Zhao, O., and *Woodroofe, M.*, 2012: Estimating a monotone trend. *Stat Sinica 22*, 359–378.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 1, January – March, 2015, pp. 69–89

Cyclic variation in the precipitation conditions of the Mátra-Bükkalja region and the development of a prognosis method

Ferenc Kovács* and **Endre Turai**

MTA-ME Research Group of Geoengineering,
Egyetem út, H-3515 Miskolc, Hungary

**Corresponding author E-mail: bgtkf@uni-miskolc.hu*

(Manuscript received in final form May 8, 2014)

Abstract—The cycle properties of the annual average, absolute maximum, and absolute minimum precipitation values have been calculated from precipitation data the Mátra and Bükk regions. The cycle parameters of annual average and annual absolute maximum precipitation values have been determined using the data of a shorter 34-year (1970–2006) and a longer 53-year (1960–2012) period (38 precipitation measurement stations) through the determination of the parameters of frequency, amplitude, and phase with an analytic version of the discrete Fourier transform (DFT), and the values obtained on the basis of the two periods have been compared. Using prognosis parameters, a prognosis until 2025 has been made. Then, the regression function of the variation in time of average and absolute maximum precipitation values has been determined on the basis of actual and prognosticated data for the whole period (1960–2025).

Key-words: Mátra-Bükkalja region, precipitation, cyclic variation, prognosis method

1. Introduction

The analysis of precipitation data in the Mátra-Bükkalja region between the years 1960 and 2012 has given the result that that both the 53-year average values of specific precipitation and the annual absolute maximum values of the measured values for the 38 precipitation measurement stations (settlements) show cylicity for both the 3–5 years and longer periods (*INNOCENTER*, 2013a; *Kovács*, 2014). Minimum and maximum 'local' values recur for both annual

average and annual maximum values. With the cyclic variation of annual precipitation values, annual average precipitation displayed constancy around the 600 mm/year value in both the Mátra and Bükkalja regions even on the basis of the combined set of data. With respect to annual absolute maximum and minimum values, regarding these parameters as indicators of extreme weather, plenty of precipitation or years of drought, the data of 53 years showed a decreasing tendency.

In the present paper, the cycle parameters of the average and absolute maximum precipitation values are calculated using the data sets reported in *INNOCENTER* (2013a) and *Kovács* (2014), analysing the precipitation data of the region investigated (Mátra-Mátraalja, Bükk-Bükkalja) and developing a calculation method of cycle parameters as a research task in the Carpathian Basin (*Szűcs*, 2012). Based on this, a prognosis is made for the period until 2025.

2. Theoretical basis of analysis and calculation, the Fourier transform

In the interpretation of frequency, amplitude. and phase, a 2π periodical $\cos(t)$ function has been taken as starting point, where $T = 2\pi$ is the period length of the function. Next, the argument of the function has been transformed (*Meskó*, 1984; *Turai*, 1983):

$$\cos(t) = \cos\left(\frac{2\pi}{2\pi}t\right) = \cos\left(\frac{2\pi}{T}t\right) = \cos\left(2\pi\frac{1}{T}t\right) = \cos(2\pi ft)$$

The rate expressing the density of periods (period density or with the commonly used term, frequency) is

$$f = \left(\frac{1}{T}\right).$$

If t stands for length in space, then frequency gives the number of periods per unit of spatial length for the given direction. Spatial frequency is called wave number.

Multiplying the $\cos(2\pi ft)$ function with factor A and shifting its maximum by Δt, after writing up

$$\cos(2\pi f[t + \Delta]), \text{ factor } A$$

is called amplitude. In the case of a monofrequency periodical signal, the amplitude equals half of the difference between the maximum (F_{max}) and minimum (F_{min}) of signal value:

$$A = \frac{F_{max} - F_{min}}{2}.$$

After a further transformation of the argument of the cosine function, the following formula can be written:

$$A\cos(2\pi f[t + \Delta t]) = A\cos(2\pi ft + 2\pi f\Delta t) = A\cos(2\pi ft + 2\pi\frac{\Delta t}{T}) = A\cos(2\pi ft + \varphi).$$

The quantity φ, thus introduced, is called phase (phase angle). the absolute phase shows the part of the phase length (phase time or wavelength) the maximum of the signal has shifted with in relation to the origin ($t = 0$). As it can be seen in *Fig. 1*, in the case of $\Delta t = 0$, the maximum shifts to the, left while in the case of $\Delta t < 0$ to the right of the origin. Absolute phase can be given in both radians and degrees:

$$\varphi = 2\pi\frac{\Delta t}{T} \quad \text{[rad]} \qquad \varphi = 360\frac{\Delta t}{T} \quad \text{[degrees]}.$$

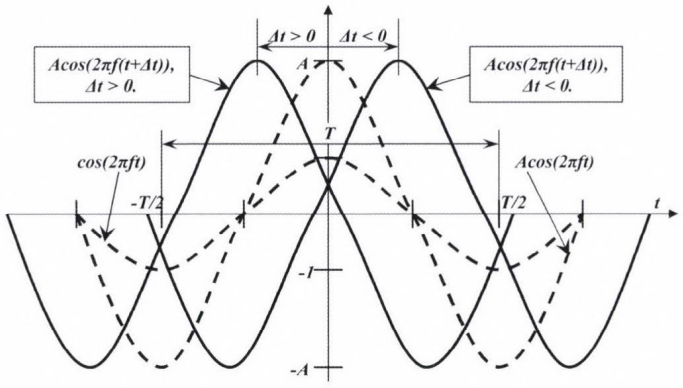

Fig. 1. The interpretation of the absolute phase.

Relative phase ($\Delta\varphi$) is interpreted between two signals and shows that in relation to the maximum of one of two signals of identical frequency, what part of the period length the maximum of the other signal has shifted with. As it can be seen in *Fig. 2*, the two signals are *x(t)* and *y(t)* while the difference of the maximums of the two signals Δt_{xy}. The relative phase between the two signals can also be calculated:

$$\Delta\varphi_{xy} = 2\pi\frac{\Delta t_{xy}}{T}\ [\text{rad}] \qquad \Delta\varphi_{xy} = 360\frac{\Delta t_{xy}}{T}\ [\text{degrees}].$$

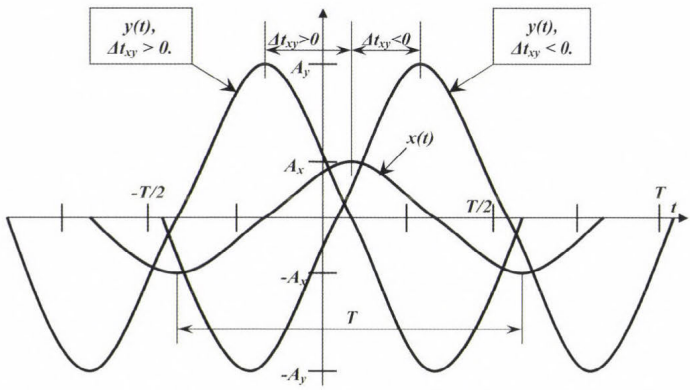

Fig. 2. The interpretation of the relative phase.

The relative phase can also be calculated as the difference of the absolute phases of the two signals:

$$\Delta\varphi_{xy} = \varphi_y - \varphi_x .$$

With the help of the Fourier transform, signals can be transferred from the space-time domain into the frequency domain. During the process, the mappings of signals in the frequency domain are called Fourier spectra.

Working with harmonic functions ($\cos(2\pi ft)$, $\sin(2\pi ft)$) in the analytic Fourier transform, a complex Fourier spectrum is obtained, which can be divided into a real and an imaginary part. The $\mathrm{Re}[F(f)]$ real part of the spectrum can be written up with a real cosine transformation

$$\mathrm{Re}[F(f)] = \int_{-\infty}^{+\infty} f(t)\cos(2\pi ft)dt , \tag{1}$$

while its imaginary part with a real sine transformation is

$$\mathrm{Im}[F(f)] = \int_{-\infty}^{+\infty} f(t)\sin(2\pi ft)dt . \tag{2}$$

The complex Fourier spectrum can be written up with two real spectra:

$$F(f) = \mathrm{Re}[F(f)] + j\,\mathrm{Im}[F(f)].$$

The real spectrum gives the weights of the cosine components falling into a frequency band unit around any f frequency, while the imaginary spectrum gives the weights of the sine components for the formation of the signal.

The $F(f)$ complex spectrum can also be defined in an exponential form by the introduction of two other real spectra:

$$F(f) = A(f)e^{j\Phi(f)}.$$

The $A(f)$ spectrum, thus introduced, is called amplitude spectrum, while the $\phi(f)$ spectrum is called phase spectrum. The amplitude spectrum gives the weight in the formation of the signal of the harmonic component falling into a frequency band unit around any f frequency, while the phase spectrum shows the part of the period length the maximum of this harmonic component shifts with in relation to the maximum of base function $\cos(2\pi ft)$, taken at point $t = 0$.

The amplitude and phase spectra are the following in the knowledge of real and imaginary spectra with the help of the correlations yielded by *Fig. 3*:

$$A(f) = \sqrt{(\mathrm{Re})^2[F(f)] + (\mathrm{Im})^2[F(f)]}$$

$$\phi(f) = arctg\frac{\mathrm{Im}[F(f)]}{\mathrm{Re}[F(f)]}$$

Real and imaginary spectrum values can also be calculated from amplitude and phase spectra:

$$\mathrm{Re}[F(f)] = A(f)\cos[\phi(f)]$$
$$\mathrm{Im}[F(f)] = A(f)\sin[\phi(f)]$$

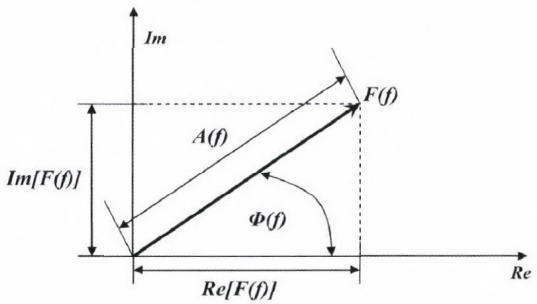

Fig. 3. Plotting of Fourier spectra in a complex plane.

3. Spectral analysis

In the search for the deterministic periodic components, the spectrum of the $\Delta y(t)$ deviations from the ($\overline{Y}$) expected values has been investigated with the following correlations:

$$\Delta y(t) = y(t) - \hat{Y},$$

$$Y(f) = \int_{-\infty}^{+\infty} \Delta y(t) e^{-j2\pi ft}\, dt.$$

The period lengths of the deterministic periodic components to be found in the stochastic signal are given by the reciprocal values of the ($f_{1,\max}$, $f_{2,\max}$, …, $f_{N,\max}$) frequencies belonging to the maximums of the $A(f)$ amplitude density spectrum of the $Y(f)$ spectrum:

$$T_1 = \frac{1}{f_{1,\max}},$$

$$T_2 = \frac{1}{f_{2,\max}},$$

…

$$T_N = \frac{1}{f_{N,\max}},$$

where N is the number of deterministic periodic components (the number of the maximums of the $A(f)$ spectrum).

It can be calculated from the $\phi(T_i)$ values of the phase-density spectrum belonging to the given period time, what $\Delta t(T_i)$ time the maximum of the given component of any T_i (i = 1, 2, …, N) period time has shifted in relation to the starting year (1973) of data registration:

$$\Delta t(T_i) = T_i \frac{\Phi(T)_i}{2\pi} \quad \text{[radian]},$$

or

$$\Delta t(T_i) = T_i \frac{\Phi(T)_i}{360} \quad \text{[degree]}.$$

The A_i amplitudes of a component of any T_i period time are given by the values of $A(f)$ amplitude density:

$$A_i = A(T_i).$$

Figure A_i gives the amplitude of the deterministic component with T_i period time.

Let $A(f)_{max}$ denote the maximum of the $A(f)$ amplitude density spectrum. The relative amplitude density spectrum normed to maximum value ($A(f)_{rel}$) as the percentage of maximum value can be calculated as follows:

$$A(f)_{rel} = \frac{A(f)}{A(f)_{max}}.100[\%].$$

Relative amplitude density spectrum values show percentage of the amplitude density of any given component of T = 1/f period time in the maximum amplitude density.

4. Spectral analysis of the variation of annual precipitation amount on the basis of Mátra Bükkalja precipitation data

In INNOCENTER (2013b) the cycle properties of the variation in time of precipitation have been investigated on the basis of the territorial average values of precipitation data in the years 1960–2012 in 23 settlements/precipitation measurement stations in the Mátra-Mátraalja region and 15 settlements/ precipitation measurement stations in the Bükk-Bükkalja region. *Table 1* shows the average annual precipitation values and the annual absolute maximum precipitation values on the basis of the data of the two regions and combined data. In order to assess the effect of the registration period on results, cycle properties have been calculated for a shorter (1973–2006, 34 years) and a longer (1960–2012, 53 years) period. (Yearbook of the Hydrographical Service of Hungary 1960–2005.)

4.1. The results of spectral analysis on the basis of precipitation data for the years 1973–2006

In the spectral analysis of the precipitation data, the registration time (T_{reg} was 33 years for end-sampling periods and the 34 years for middle-sampling periods. The sampling rate (Δt) was 1 year, while the number of samples was 34.

Analyses have been performed with an analytic version of the discrete Fourier transform (DFT) (*Turai*, 1983). The complex amplitude density spectra of the function of annual precipitation values have been determined as the function of discrete period time values. Of the four real spectra describing the complex spectrum (real spectrum, imaginary spectrum, amplitude spectrum, and phase spectrum), amplitude spectra are presented. In the plotting, logarithmic linear scale has been chosen to illustrate spectrum maximums more clearly.

Table 1. Precipitation data in the Mátra-Bükkalja region as the function of time

Number of settlements / measurement points		Years												
		1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972
Mátra	Average	727	506	568	681	747	868	814	538	478	737	861	455	636
	Absolute maximum	1016	584	724	923	1012	1098	1029	610	599	899	1080	559	789
Bükk	Average	712	436	499	548	693	825	769	569	517	623	891	378	557
	Absolute maximum	811	487	535	599	769	947	956	733	563	764	971	410	674
Mátra Bükk	Average	721	476	538	624	724	849	795	551	495	688	874	422	602
	Absolute maximum	1016	584	724	923	1012	1098	1029	733	599	899	1080	559	789

Number of settlements / measurement points		Years												
		1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986
Mátra	Average	509	804	607	709	620	576	691	659	504	444	425	655	674	453
	Absolute maximum	712	1061	748	935	748	704	844	821	595	555	504	791	814	566
Bükk	Average	453	698	675	659	691	585	652	697	524	423	504	535	594	399
	Absolute maximum	535	735	843	747	781	708	763	740	574	506	581	648	680	438
Mátra Bükk	Average	485	759	636	688	650	580	674	676	512	435	459	604	640	430
	Absolute maximum	712	1061	843	935	781	708	844	821	595	555	581	791	814	566

Number of settlements / measurement points		Years													
		1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Mátra	Average	534	638	611	560	692	393	514	497	718	637	433	654	949	474
	Absolute maximum	654	851	729	613	774	486	603	586	874	751	519	754	1092	563
Bükk	Average	499	596	586	492	652	346	457	454	679	576	482	722	874	426
	Absolute maximum	559	613	655	523	694	418	516	509	727	609	535	774	988	474
Mátra Bükk	Average	519	620	600	531	675	373	490	479	701	611	454	683	917	453
	Absolute maximum	654	851	729	613	774	486	603	586	874	751	535	774	1092	563

Number of settlements / measurement points		Years												
		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	
Mátra	Average	585	639	551	647	751	651	632	657	666	1054	404	434	
	Absolute maximum	723	837	674	747	929	733	791	777	736	1195	462	486	
Bükk	Average	653	591	517	749	741	631	599	625	624	1118	447	466	
	Absolute maximum	736	669	564	828	791	711	649	701	709	1153	488	557	
Mátra Bükk	Average	614	618	536	690	746	642	619	645	658	1079	420	446	
	Absolute maximum	736	837	674	828	929	733	791	777	736	1195	488	557	

In the calculation of spectra, the spectrum of $\Delta y(t)$ deviations from ($\bar{Y}$ $\hat{Y}$) expected values has been determined:

$$\Delta y(t) = y(t) - \bar{Y}.$$

The (T_{min}) minimum period time that can theoretically be found in the signal is defined by the Nyquist frequency (f_N).

$$\Delta t = 1 \text{ year}, \qquad f_N = 0{,}5 \frac{1}{year}, \qquad T_{min} = 2 \text{ years}.$$

As in the case of all the six time series, the 'sampling' time was 1 year, the analysis can only reveal cycles of longer period time than 2 years in the changes everywhere.

In theory, maximum period time (T_{max}) is determined by the registration time (T_{reg}):

$T_{max} = T_{reg}$ – in case of end sampling,

$T_{max} = T_{reg} + \Delta t$ – in case of middle sampling.

Therefore, the maximum period time that can be revealed by analysis is 33 years in case of end-sampling and 34 years in case of middle-sampling

With the data in *Table 1*, both the amplitude spectra of the amplitude density and the relative spectra have been determined. In the latter case, spectra have been normed to maximum spectrum value. In all the six cases – annual average and annual absolute maximum precipitation, – for Mátra, Bükk, and Mátra+Bükk regions, similar amplitude and relative amplitude spectrum functions have been obtained.

The cycle properties of annual average precipitation in the Mátra region are the following on the basis of amplitude peaks, cycle time, and amplitude density:

Major cycles: 1. T_1 = 4.9 years, A_1 = 1243 mm; 2. T_2 = 3.5 years, A_2 = 1195 mm; 3. T_3 = 29.8 years, A_3 = 946 mm; 4. T_4 = 9.9 years, A_4 = 806 mm; minor cycles: 1. T_1 =7.3 years, A_1 = 476 mm; 2. T_2 = 6.3 years, A_2 = 440 mm.

Cycle properties revealed on the basis of Bükk data are, cycle time and amplitude density: major cycles: 1. T_1 = 28.7 years, A_1 = 1216 mm; 2. T_2 = 3.5 years, A_2 = 1064 mm; 3. T_3 = 4.9 years, A_3 = 1035 mm; 4. T_4 = 9.5 years, A_4= 929 mm; minor cycles: 1. T_1 = 7.3 years, A_1 = 541 mm; 2. T_2 = 6.1 years, A_2 = 308 mm.

The combined treatment of Mátra+Bükk data has also revealed 4 major and 2 minor cycles in the variation of annual precipitation values (*Figs. 4* and *5*),

cycle time and amplitude density: major cycles 1. T_1 = 5.0 years, A_1 = 1,139 mm; 2. T_2 = 3.5 years, A_2 = =1,119 mm; 3. T_3 = 29.2 years, A_3 = 1,080 mm; 4. T_4 = 9.7 years, A_4 = 860 mm; minor cycles 1. T_1 = 7.4 years, A_1 = 508 mm; 2. T_2 = 6.2 years, A_2 = 310 mm.

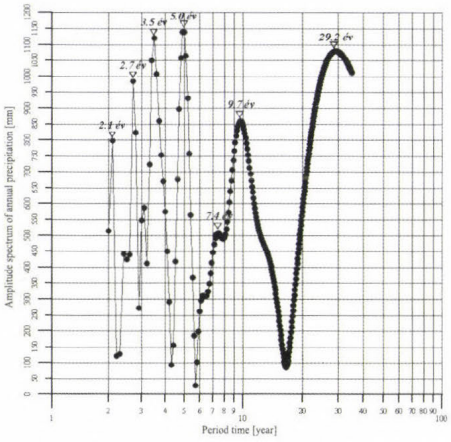

Fig. 4. Amplitude spectrum of annual precipitation int he Mátraalja and Bükkalja regions.(sampling rate = 1 year)

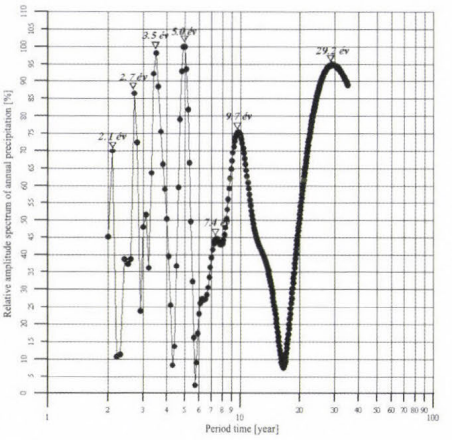

Fig. 5. Relative amplitude spectrum of annual precipitation in the Mátraalja and Bükkalja regions. (sampling rate = 1 year)

Cycle properties that can be revealed on the basis of the amplitude spectrum and relative amplitude spectrum detected in the variation of annual absolute maximum precipitation values, cycle time, and amplitude density for the Mátra region are the following: 1. T_1 = 3.5 years, A_1 = 1561 mm; 2. T_2 = 5.0 years, A_2 = 1434 mm; 3. T_3 = 10.9 years, A_3 = 1352 mm; 4. T_4 = 31.4 years, A_4 = 1262 mm; minor cycles 1. T_1 = 7.5 years, A_1 = 741 mm; 2. T_2 = 6.2 years, A_2 = 474 mm.

Cycle properties of the Bükk region are: major cycles 1. T_1 = 27.0 years, A_1 = 1408 mm; 2. T_2 = 3.4 years, A_2 = 1297 mm; 3. T_3 = 5.0 years, A_3 = 1168 mm; 4. T_4 = 9.7 years, A_4 = 973 mm; minor cycles 1. T_1 = 7.4 years, A_1 = 796 mm; 2. T_2 = 6.2 years, A_2 = 362 mm.

Cycle properties of Mátra and Bükk combined data on the basis of amplitude spectra (*Figs. 6* and *7*), cycle time, and amplitude density are the following: major cycles 1. T_1 = 3.5 years, A_1 = 1482 mm; 2. T_2 = 5.0 years, A_2 = 1413 mm; 3. T_3 = 30.3 years, A_3 = 1256 mm; 4. T_4 =11.1 years, A_4 = 1225 mm; minor cycles 1. T_1 = 7.5 years, A_1 = 734 mm; 2. T_2 = 6.2 years, A_2 = 298 mm.

On the basis of the above results, the following generalizations can be made:

— In the case of the six time series examined with respect to annual precipitation variation, cycles of approximately identical period times can be revealed.

— In the case of all the six time series, there have been found periods of 3.5 years, 5 years, 10–11 years, and 27–31 years as major cycles.
— In all the cases, 6.2-year and 7.3–7.5-year periods appear as minor cycles. (To prove the existence of 27–31-year cycles in a more reliable way, longer data series would be needed.)

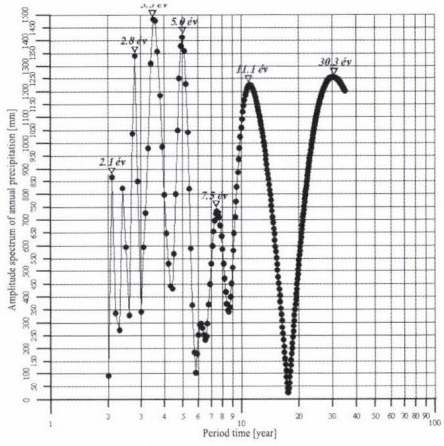

Fig. 6. Amplitude spectrum of annual precipitation in the Mátraalja and Bükkalja regions (sampling rate = 1 year).

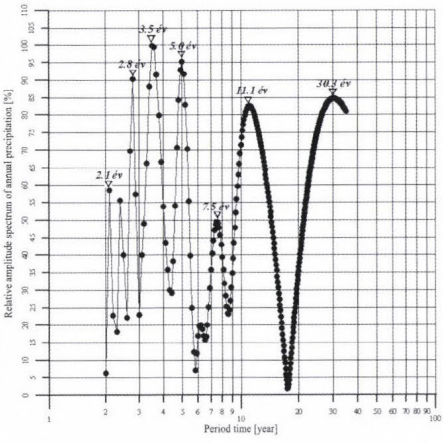

Fig. 7. Relative amplitude spectrum of the absolute maximum of annual precipitation in the Mátraalja and Bükkalja regions (sampling rate = 1 year.)

4.2. The results of spectral analysis on the basis of precipitation data in the years 1960-2012

The registration period is 1960–2012, the length of the registration period (T_{reg}) is 52 years with end sampling and 53 years with middle sampling, sampling rate (Δt) is 1 year, the number of samples is 53. The calculation process has been according to Section 4.1, the maximum period time that the analysis can reveal is T_{max} = 52 years – 53 years.

On the basis of amplitude peaks, the following precipitation cycles can be revealed for the Mátra annual precipitation values, cycle time, and amplitude density: major cycles T_1 = 5.0 years, A_1=2765 mm; 2. T_2 = 3.6 years, A_2 = 2074 mm; 3. T_3 = 41.1 years, A_3 = 1555 mm; 4. T_4 = 10.7 years, A_4 = 1494 mm; minor cycles 1. T_1 = 6.4 years, A_1 = 1101 mm; 2. T_2 = 5.7 years, A_2 = 1027 mm; 3. T_3 = 8.6 years, A_3 = 675 mm; 4. T_4 = 14.3 years, A_4 = =642 mm; 5. T_5 = 7.4 years, A_5 = 577 mm; 6. T_6 = 19.8 years, A_6 = 456 mm.

In the Bükkalja region, the following cycles can be revealed in the variation of annual precipitation values on the basis of amplitude spectrum and relative amplitude spectrum, cycle time, and amplitude density: major cycles 1. T_1 =

5.0 years, A_1 = 2567 mm; 2. T_2 = 38.6 years, A_2 = 1759 mm; 3. T_3 = 10.5 years, A_3 = 1747 mm; 4. T_4 = 3.6 years, A_4 = 1719 mm; minor cycles 1. T_1 = 5.7 years, A_1 = 1413 mm; 2. T_2 = 6.5 years, A_2 = 1220 mm; 3. T_3 = =14.2 years, A_3 = 753 mm; 4. T_4 = 7.5 years, A_4 = 552 mm; 5. T_5 = 8.4 years, A_5 = 504 mm; 6. T_6 = 19.8 years, A_6 = 323 mm.

The combined treatment of Mátra and Bükk data also reveals 4 major and 6 minor cycles on the basis of annual precipitation values (*Figs. 8* and *9*), cycle time and amplitude density: major cycles: 1. T_1 = 5.0 years, A_1 = 2685 mm; 2. T_2 = 3.6 years, A_2 = 1928 mm; 3. T_3 = 40.4 years, A_3 = 1635 mm; 4. T_4 = 10.6 years, A_4 = 1587 mm; minor cycles 1. T_1 =5.7 years, A_1 = 1188 mm; 2. T_2 = 6.4 years, A_2 = 1151 mm; 3. T_3 = 14.2 years, A_3 = 669 mm,;4. T_4 = 8.5 years, A_4 = 592 mm; 5. T_5 = 7.4 years, A_5 = 577 mm; 6. T_6 = 20.0 years, A_6 = 383 mm.

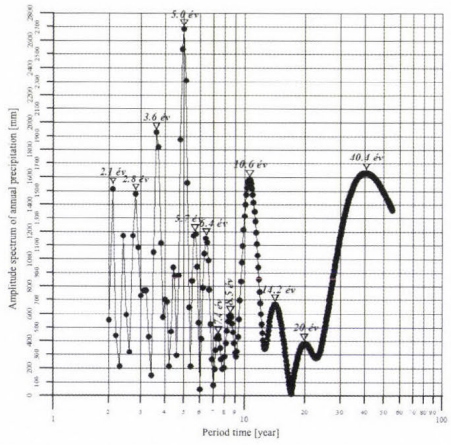

Fig. 8. Amplitude spectrum of annual precipitation in the Mátraalja and Bükkalja regions (sampling rate = 1 year).

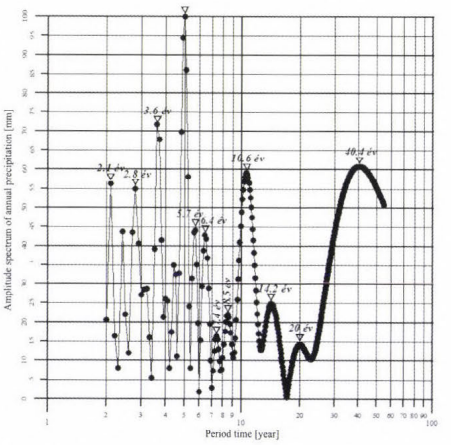

Fig. 9. Relative amplitude spectrum of annual precipitation in the Mátraalja and Bükkalja regions (sampling rate = 1 year).

The analysis of the absolute maximum values of annual precipitation reveals the following cycle properties on the basis of Mátra data, cycle time, and amplitude density: major cycles 1. T_1 = 5.0 years, A_1 = 3306 mm; 2. T_2 = 3.6 years, A_2 = 2656 mm; 3. T_3 = 45.6 years, A_3 = 2119 mm; 4. T_4 = 10.8 years, A_4 = 1806 mm; minor cycles 1. T_1 = 5.6 years, A_1 = 1319 mm; 2. T_2 = 6.4 years, A_2 = 1191 mm; 3. T_3 = 13.9 years, A_3 = 1044 mm; 4. T_4 = 7.3 years, A_4 = 1046 mm; 5. T_5 = 8.6 years, A_5 = 814 mm; 6. T_6 = 19.8 years, A_6 = 722 mm.

Similarly, 4 major and 6 minor cycles can be revealed on the basis of the Bükk-Bükkalja absolute maximum precipitation data, cycle time and amplitude density: major cycles 1. T_1 = 5.0 years, A_1 = 2646 mm; 2. T_2 = 38.6 years, A_2 = 2138 mm; 3. T_3 = 10.5 years, A_3 = 2024 mm; 4. T_4 = 3.6 years, A_4 = 1758 mm; minor cycles 1. T_1 = 5.6 years, A_1 = 1434 mm; 2. T_2 = 6.4 years, A_2 = 1351 mm;

3. T_3 = 14.0 years, A_3 = 885 mm; 4. T_4 = 8.4 years, A_4 = 883 mm; 5. T_5 = 7.3 years, A_5 = 445 mm; 6. T_6 = 19.4 years, A_6 = 454 mm.

Cycle properties of absolute maximum precipitation values cycle time, and amplitude density in the combined assessment of the Mátra+Bükk region (*Figs. 10* and *11*) are the following: major cycles 1. T_1 = 5.0 years, A_1 = 3,168 mm; 2. T_2 = 3.6 years, A_2 = 2468 mm; 3. T_3 = 46 years, A_3 = 2271 mm; 4. T_4 = 10.7 years, A_4 = 1842 mm; minor cycles 1. T_1 = 5.7 years, A_1 = 1273 mm; 2. T_2 = 6.4 years, A_2 = 1127 mm; 3. T_3 = 13.7 years, A_3 = 982 mm; 4. T_4 = 8.5 years, A_4 = 721 mm; 5. T_5 = 7.3 years, A_5 = 726 mm; 6. T_6 = 19.6 years, A_6 = 714 mm.

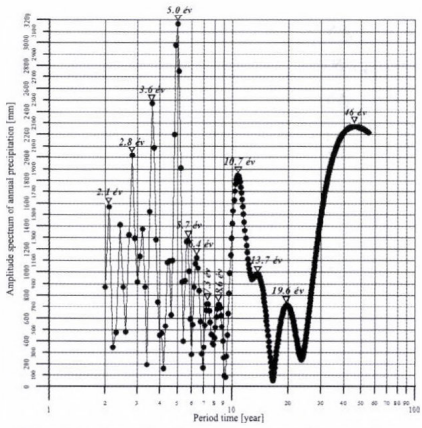

Fig. 10. Amplitude spectrum of annual precipitation in the Mátraalja and Bükkalja regions (sampling rate = 1 year) .

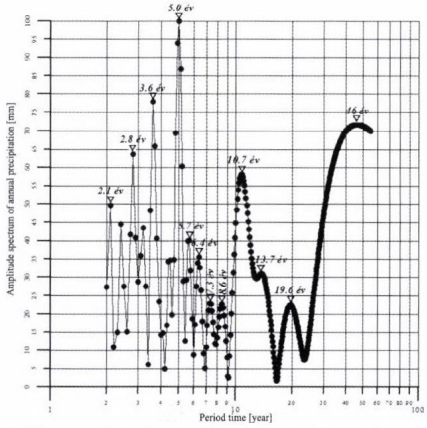

Fig. 11. Relative amplitude spectrum of the absolute maximum of annual precipitation in the Mátraalja and Bükkalja regions (sampling rate = 1 year) .

From the cycle properties determined on the basis of the data of precipitation time series of 53 years, the following generalizations can be made:

— Cycles of nearly identical period time can be revealed on the basis of the six time series investigated with respect to annual precipitation variation.

— In the case of all the six time series, the 3.6-year, the 5-year, the 10.5–10.8-year and 38.6–46-year periods appear as major cycles.

— In all the cases, the 5.6–6.7-year, 6.4-year, 7.3–7.5-year, 8.4–8.6-year, 13.7–14.3-year and 19.4–20-year periods appear as minor cycles.

The comparison of the cycle time data of the major and minor cycles revealed on the basis of the two time series of different lengths (34 years and 53 years) has yielded the following results:

— With all the three data groups, the number of major cycles that can be revealed on the basis of both time series is the same: four.

— In the case of the shorter time series, 2 minor cycles have been found for all the three data groups, while for the longer time series (53 years), 6 minor cycles have been revealed.

— With the shorter, generally maximum 10-year cycle times, practically identical/equivalent cycle time has been revealed for both the major and minor cycles, namely, in Mátra: 3.5–3.6 years, 4.9–5.0 years, 9.9–10.7 years, 6.3–6.4 years, 7.3–7.4 years,in Bükk: 3.5–3.6 years, 4.9–5.0 years, 9.5–10.5 years, in Mátra+Bükk: 3.5–3.6 years, 5.0–5.0 years, 9.7–10.6 years, 6.2–6.4 years, 7.4–7.4 years.

— In all the three areas, it has been identically found for longer cycle times (above 30 years) that on the basis of the 34-year time series, a shorter major cycle time, while on the basis of longer time series, a longer major cycle time has been revealed, namely, in Mátra: 29.8 years, 41.1 years, in Bükk: 28.7 years, 38.6 years, in Mátra+Bükk: 29.2 years, 40.4 years.

The differences found in the latter case confirm the former observation that for a long-time prognosis, a time (data) series longer than 50 years is required.

5. Determination of prognosis values

On the basis of Sections 2 and 3, including the summary of the basics of spectral data processing, the $y(t)$ time series of precipitation values can be restored through the 'use' of the $A(f)$ amplitude density and $\phi(f)$ phase density spectra, defined in the previous analyses:

$$y(t) = \overline{Y} + \int_{-f_N}^{+f_N} A(f) e^{j[2\pi ft + \phi(f)]} df ,$$

where f_N is the Nyquist frequency and it equals to 0.5 year^{-1}.

As the Fourier spectrum is even, the former equation can also be written up in the following form:

$$y(t) = \overline{Y} + \int_{o}^{+f_N} A(f) e^{j[2\pi ft + \phi(f)]} df .$$

With the use of the T_i (i = 1,2,..., N=10) period times of major and minor cycles, the A_i (i = 1,2,..., N = 10) amplitude, and the $\phi(T_i)$ (i = 1,2,..., N = 10) phase values, it is possible to define the $[y(t)^{det}]$ time series of the amount of precipitation attributable to deterministic causes:

$$y(t)^{\det} = \overline{Y} + \frac{2}{T_{reg}} \sum_{i=1}^{10} A_i \cos\left[\frac{2\pi}{T_i}(t-1960) + \phi(T_i)\right].$$

Using the $\{R_e[F(T_i)]\}$ and $\{I_m[F(T)_i]\}$ values calculated for given T_i period times of real and imaginary spectra, the $\phi(T_i)$ phases of the specific components can be defined with the following correlation:

$$\phi(T_i) = arctg \frac{I_m[F(T_i)]}{R_e[F(T_i)]}.$$

The difference between the $y(t)$ actual time series and $y(t)^{det}$ represents the accidental (stochastic) impact.

If $t > 2012$ values are put in the former equation, the amount of precipitation that can be expected in the given years can be estimated (forecast) with extrapolation. It must be added, however, that this estimation would only yield a prognosis of 100% reliability by using spectra calculated from an infinitely large $y(t)$ registratum (annual data), which, of course, cannot be expected in the case of the 53 years long time series investigated.

Furthermore, there is a possibility of estimating periodicity with modern statistical methods (analysis with autocorrelation functions, factor and cluster analysis), although these tools would only give similarly precise results as the spectral analysis applied on the basis of data series of several hundred years.

Using the spectrum data in *Fig. 8*, taking into account the impact of the four deterministic major cycles (5, 3.6, 40.4, and 10.6 years) and taking into consideration the impact of the further 6 minor cycles in *Fig. 12* as well as that of the two cycles (2.1 years and 2.8 years) earlier omitted due to aliasing distortion, the prognosis values in *Figs. 13, 14* are obtained. According to *Fig. 4*, the two short cycles are present in the prognosis of annual precipitation values with a relatively high amplitude, above 55%, there has been a spectacular improvement in classic statistical indicators. Deviation (RMS) has decreased from 16.1% and 15.7% to 12.6%, while the correlation coefficient (r) has increased from 0.78 and 0.79 to 0.89.

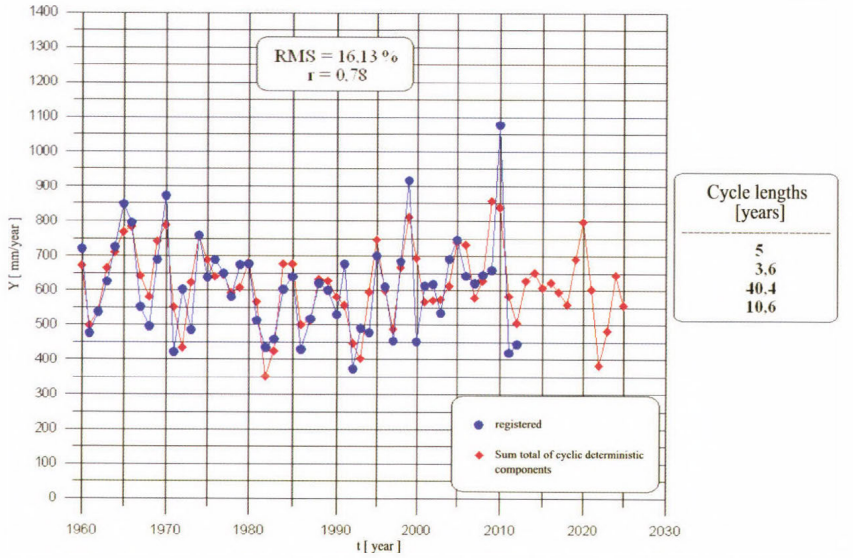

Fig. 12. Annual precipitation value in the Mátraalja and Bükkalja regions. (Prognosticated on the basis of four cyclic components.)

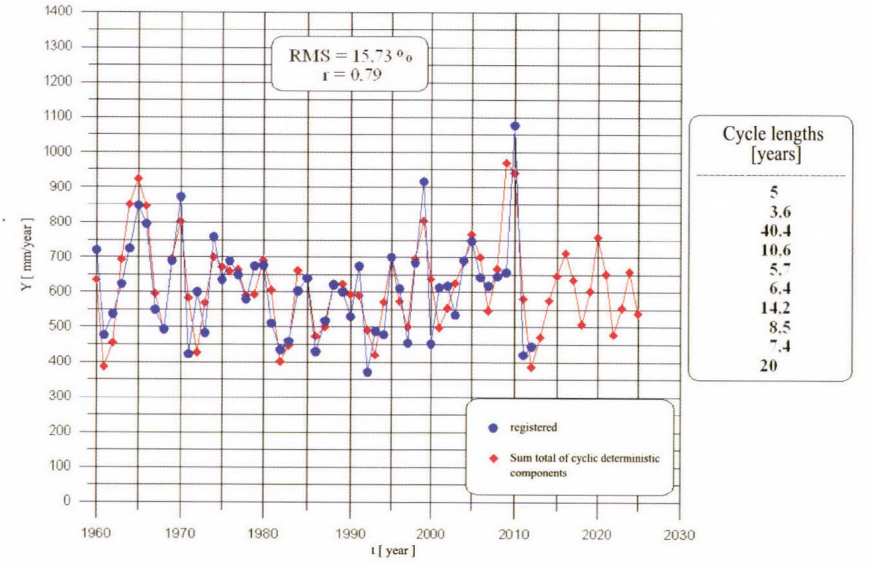

Fig. 13. Annual precipitation value in the Mátraalja and Bükkalja regions. (Prognosticated on the basis of ten cyclic components.)

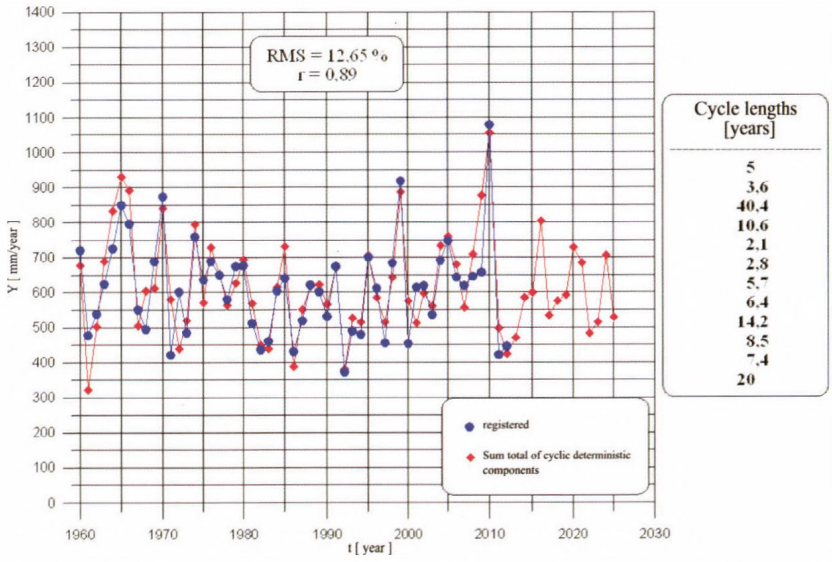

Fig. 14. Annual precipitation value in the Mátraalja and Bükkalja regions. (Prognosticated on the basis of twelve cyclic components.)

The amplitude data in *Fig. 10* and relative amplitude data in *Fig. 11* have been used in the calculation of annual absolute maximum precipitation prognosis. Taking the four deterministic and the further 6+2 cycle properties into account, the absolute maximum precipitation prognosis in *Figs. 15, 16,* and *17* has been obtained.

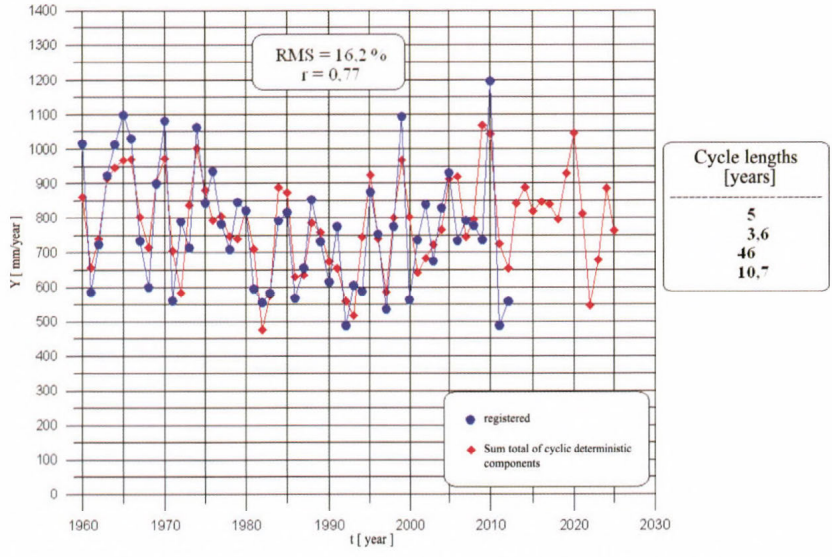

Fig. 15. Variation in the annual maximum of annual precipitation value in the Mátraalja and Bükkalja regions. (Prognosticated on the basis of four cyclic components.)

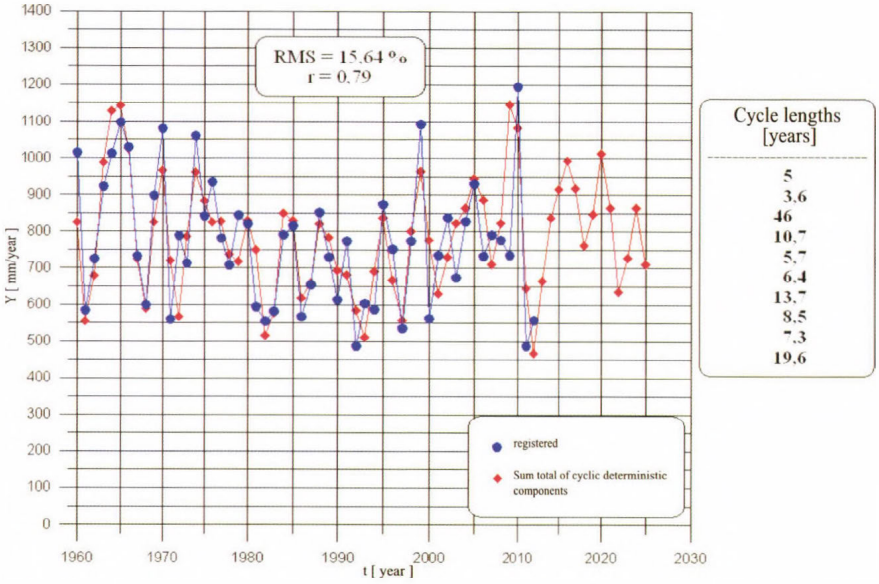

Fig. 16. Variation in the annual maximum of annual precipitation value in the Mátraalja and Bükkalja regions (Prognosticated on the basis of ten cyclic components.)

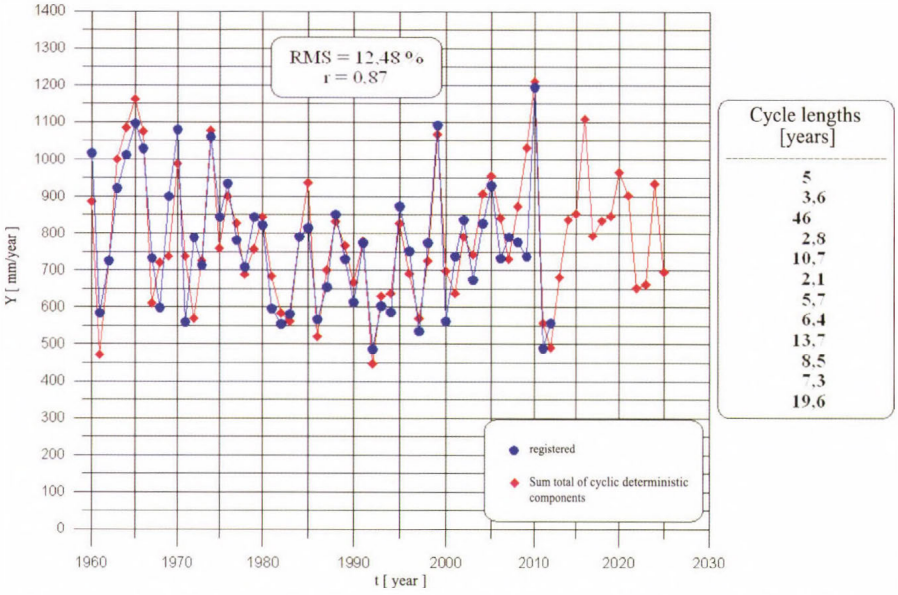

Fig. 17. Variation in the annual maximum of annual precipitation value in the Mátraalja and Bükkalja regions. (Prognosticated on the basis of twelve cyclic components.)

On the basis of the classical statistical parameters (RMS = 16.2%, r = 0.77) it can be concluded here, too, that between 1960 and 2012, the four deterministic major cycles decisively determined absolute maximums (*Fig. 15*). Taking the six minor cycles into account hardly improves classical statistical parameters (RMS = 15.6 %, r = 0.79) in this case, either, but the prognosticated sections in *Figs. 15* and *16* are significantly different here, too. Taking into account the two short cycles (2.1 years and 2.8 years), also appearing here with a high amplitude, has considerably improved classical statistical indicators (RMS = 12.5%, r = 0.87) (*Fig. 17*).

On the basis of the data in *Fig. 14*, for the purpose of practical utilisation it can be underlined in the prognosis, that the exceedingly high, 1079 mm/year amount of precipitation of 2010 – a uniquely high value in the last 53 years – will not recur in the next 12–15 years. The 850–900 mm/year annual precipitation, having occurred several times in previous years (1965, 1970, 1999) may 'probably be expected' in 2016. On the other hand, it is good news that in the coming 12–15 years, no annual precipitation below 500 mm/year, causing severe drought, may be expected.

The 1100 mm/year maximum precipitation prognosticated for 2016 (see *Fig. 17*), remains 100 mm/year below the round 1200(1195) mm/year value of 2010 but may reach the 1100 mm peak data of the years 1965, 1970, 1974, and 1999.

6. Variation in time of precipitation properties between the years 1960 and 2025

With the combined handling of the actual data for the years 1960–2012, presented in *Table 1* and *Figs. 14* and *17*, and the prognosis data in *Figs. 14* and *17* related to the Mátra+Bükk region, the time function of the variation of annual precipitation, and the absolute maximum precipitation values for the years 1960–2025 have been determined with the conventional statistical method.

The function in *Fig. 18* shows a constancy of 620–605 mm/year of annual (average) precipitation with 0.23 = 23% empirical deviation ($D_{deg}/Y_{average}$). The correlation coefficient characterizing the closeness of the function determined from the data of the 65-year time series is r^2 = 0.00048, which indicates the independence of the two variables of annual precipitation (average) and time (years) according to conventional statistical interpretation.

Fig. 19 shows the regression function determined on the basis of actual and prognosed annual absolute maximum precipitation data between the years 1960 and 2025. With an acceptable (reliable) 19% corrected empirical deviation and a r^2 = 0,00027 regression coefficient, the function predicts the constancy of the annual absolute maximum in the statistical sense while, for example, it predicts a 1100 mm precipitation maximum for 2016.

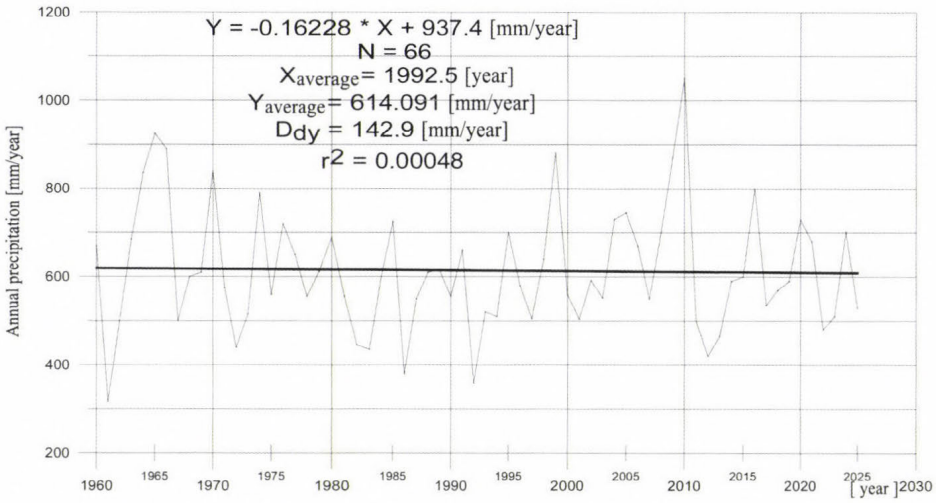

Fig. 18. Regression function of the variation in time of the annual precipitation conditions (1960-2012) and prognosis data (2013-2025) of the Mátra-Bükkalja region.

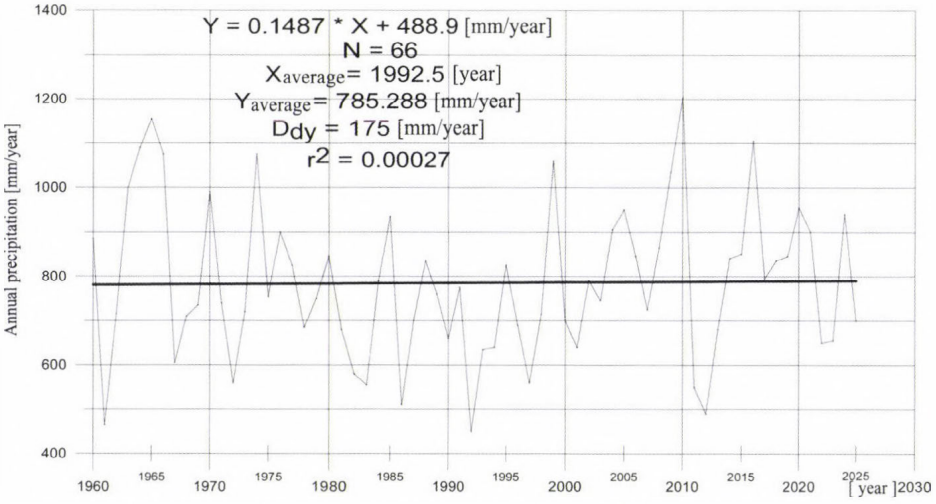

Fig. 19. Regression function of the variation in time of the annual absolute maximum precipitation properties (1960–2012) and prognosis data (2013–2025) of the Mátra-Bükkalja region.

Acknowledgements: The research was carried out in the framework of the Sustainable Resource Management Centre of Excellence at the University of Miskolc, as part of the TÁMOP-4.2.2/A-11/1-KONV-2012-0049 „WELL aHead" project in the framework of the New Széchenyi Plan, funded by the European Union and co-financed by the European Social Fund.

References

INNOCENTER, 2013a: A Mátra-Bükkalja csapadék jellemzőinek meghatározása, a csapadék jellemzők időbeli alakulása. Kutatási részjelentés a Mátrai Erőmű Zrt. részére, 2013. augusztus. (in Hungarian)

INNOCENTER, 2013b: A Mátra-Bükkalja csapadék jellemzői ciklikus változása, prognózis módszer kidolgozása. Kutatási részjelentés a Mátrai Erőmű Zrt. részére, 2013. augusztus. (in Hungarian)

Kovács, F., 2014: A Mátra-Bükk-i terület csapadék jellemzőinek alakulása az utóbbi ötven évben (1960–2012). *Bányászati és kohászati lapok. Bányászat 147*, 2–6 (in Hungarian)

Meskó, A., 1984: Digital Filtering. Akadémiai Kiadó, Budapest

Szűcs, P., 2012: Hidrogelógia a Kárpát-medencében – hogyan tovább? *Magyar Tudomány 173*, 554–565. (in Hungarian)

Turai, E., 1983: A Fourier transzformáció egy numerikus módszere és alkalmazása a GP-jelenség rendszerelemző függvénnyel történő leírásánál. *Magyar Geofizika XXIV*, 11–19.

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 1, January – March, 2015, pp. 91–109

Analyzing long-term evapotranspiration of Lake Fenéki wetland (Kis-Balaton, Hungary) between 1970 and 2012

Angéla Anda[1*], Katalin Nagy[2], Gábor Soós[1], Tamás Kucserka [1]

[1]Pannon University Georgikon Faculty,
Department of Meteorology and Water Management
Festetics u. 7, H-8360 Keszthely, Hungary

[2]West-transdanubian Water Inspectorate,
Vörösmarty u. 2, H-9700 Szombathely, Hungary

**Corresponding author E-mail: anda-a@georgikon.hu; anda@keszthelynet.hu*

(Manuscript received in final form March 12, 2014)

Abstract—The aim of the study was to estimate long term evapotranspiration (ET) of Kis-Balaton wetland through the investigation of Lake Fenéki. Data set was processed using the West-transdanubian Water Inspectorate methodology. Potential evapotranspiration (PET) was calculated using Hungarian empirical models (*Antal* and *Dunay*), while Lake Fertő formula was applied evaluating the ET that includes the impacts of vegetation.

Calculated PET values of the wider (adjacent) environment of Lake Fenéki (Zalaegerszeg, Nagykanizsa, and Keszthely meteorological stations) differed significantly and further variation was observed in PET, when measured meteorological elements on Lake Fenéki were applied. PET increment, as a result of linear trend fitted to the 43-year long data (Keszthely station) was 3–4 mm year^{-1}. Relation between PET calculated from the data of Keszthely station and for Lake Fenéki was strong, so PET of Lake Fenéki can be originated from the data of Keszthely station. Calculated ET was not significantly different due to the likely similar input data in ET calculation model of Hungarian Meteorological Service (OMSZ).

43-year annual mean ET for Lake Fenéki was 809±88 mm. This ET was 84% of calculated PET. Analyzing the nine dry-warm seasons, average annual ET exceeded the long term average (874.7±37.6 mm) with 78 mm. The average ET of the remaining 34 wet-cold seasons totalled 796.6±89.4 mm.

Empirical formulas cannot be replaced, according to monthly ET comparisons, by using "A" class pan estimating the ET of aquatic habitats.

Seasonal pattern of monthly ET time series for Lake Fenéki was analyzed using autoregressive integrated moving averages (ARIMA) modeling technique. After first differencing, the transformed series was stationary and found to be governed by moving average process of order 1.

Key-words: wetland evaporation, potential evapotranspiration, seasonal ARIMA, Kis-Balaton wetland (Lake Fenéki)

1. Introduction

Kis-Balaton is a large, continuous wetland. It used to belong to Lake Balaton (Western bay). In 1863, suitable circumstances turned up to control the water level more or less separately from meteorological conditions (*Harkay*, 1983; *Virág*, 1997). Prior to the opening of the Sió-sluice in 1863, the water level of Lake Balaton was determined by the prevailing weather conditions. When the sluice gate of Channel Sió was built, the natural fluctuation of water level decreased to 0.5 m or less (earlier the fluctuation of water level could be as high as 3–5 meters). In the first part of the 19th century, the water level of Lake Balaton was lowered in several steps, therefore, the higher areas of the Kis-Balaton basin dried out. The water level turned to 2–3 m lower than the water level of the ancient lake (*Kovács et al.*, 2010; *Hatvani et al.*, 2011). The surface of Kis-Balaton wetland became smaller, water cover of higher parts disappeared, and it appeared again when River Zala flooded. In the 1920's banks were built to both sides of River Zala. As a consequence of the regulation, the Kis-Balaton and other surrounding marshes lost their function to protect the water quality of Lake Balaton. It was strengthened by the consequences of civilization, i.e., intensive agricultural chemicalization, increasing urbanization, developing and spreading of holiday resorts in the region, which altogether resulted in the significant deterioration of quality of waters entering Lake Balaton. This marsh supplied an ideal habitat for plants and animals preferring aquatic habitats. As the area became drier, the surface of the marsh got smaller (*Nguyen et al.*, 2005). At the deepest part – close to Island Diás –, two small lakes and the surrounding reed (about 1400 ha) remained behind. This area is under strict protection according to *Ramsar Convention* (1971). Because of the regulation works, the water protecting function of Kis-Balaton and the ambient groves ended.

In the 1960–70's, especially in Keszthely Basin, water quality of Lake Balaton notably decayed (*Istvánovics et al,.* 1997; *Padisak* and *Reynolds*, 2003), so a resolve was made about the artificial reconstruction of the marsh (Kis-Balaton Water Protection System, KBWPS) and the restitution of the water protecting function of Kis-Balaton (*Pomogyi et al.*, 1996; *Tátrai et al.*, 2000). Construction works began in 1981, the Stage 1, called Lake Hídvég, was constructed between 1981 and 1985 (*Korponai et al.*, 2010).

Due to economic reasons, the construction did not follow the planned timing, so in behalf of water quality of Lake Balaton a makeshift was made. In 1992, the north-northwest part of Lake Fenéki was flooded. The area was called Grove Ingói (16 km^2).

Constructions were made at the non-flooded area of Lake Fenéki as well (outer reservoir). The water of Zala-Somogy border ditch and the Marótvölgyi Channel was led to this area. One part of the effluent water from Lake Hídvégi – mostly in case of fill of water – was directed to the non-flooded area. This

caused a temporary flood and reservation. Complete construction of the Kis-Balaton Water Protection System Stage 2 is still in process (Project KEOP 2.2.1/2F/09-2009-0001). In 2015, the flood of outer reservoir is expected to happen. Coordination of water quality protection, conservation, flood and water conservation objectives and tasks is going to occur in the project.

In case of flow-through lakes (like Kis-Balaton), inflow and outflow are the most important members of the water balance. In Lake Fenéki, this value can be as high as 90% on annual basis. Although the share of evaporation is usually just under 10%, accurate estimation of evapotranspiration (*ET*) is important, because during summer – in the lack of notable flood – it can reach 30–70% of inflow and outflow. Typically, in case of low water conditions, it is problematic prescribing the water balance.

Evaporation is a major energy-consuming physical phenomenon, which can be significantly affected by climate change (*Novaky*, 2005). To refine the hydrological and climatic forecasts, it is necessary to estimate evaporation irrespectively of the other members of water balance equation.

There are only a very few studies which determine *ET* in the area of Kis-Balaton. Hungarian researchers have investigated only the *ET* of bigger lakes (*Havalda,* 1930; *Szesztay*, 1962; *Antal,* 1968; *Antal et al.,* 1977; *Varga,* 2005; *Varga et al.,* 2007).

This research aims to overview the practice of *ET* calculation used by the Inspectorate for Water Management for decades, and to produce a long time series *ET* dataset for Lake Fenéki. Time series is a collection of quantitatively measured data evenly spaced in time. Analysis of the time series is important to understand the structure and functioning of observed data. Time series analysis allows a mathematical model to be built in order to discuss the data trend (*Box et al.,* 2008; *Brockwell* and *Davis*, 2001; *Psilovikos* and *Elhag*, 2013). The autocorrelation elements of the data were extracted using ARIMA to model the underlying wetland evaporation trend more precisely.

2. Materials and methods

Our study aimed ET of Lake Fenéki (46°38'N, 17°11'E, altitude 194 m, area 54–57 km^2), which is situated in the natural valley of River Zala between Balatonhídvég and the mouth of Zala, and it is the planned reservoir of the Kis-Balaton Water Protection System Stage 2. On the northern part of the reservoir, water is kept by the southern watertight dam, on the western part by the Zalavár dam. On the east side, Zala valley is closed 3.5 km wide by the valley occlusion parallel with the railway line (*Fig. 1*).

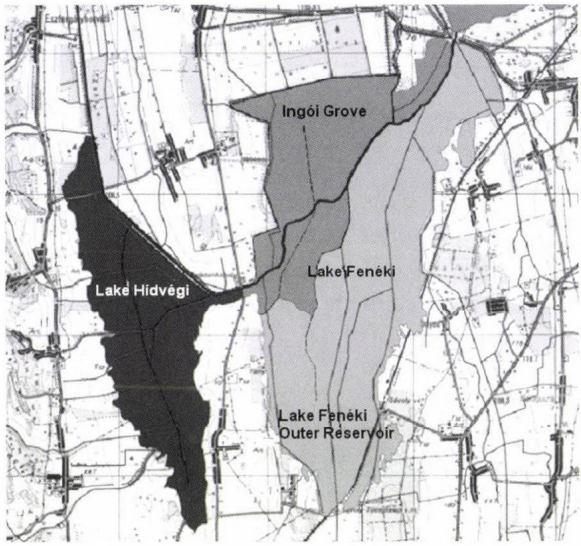

Fig. 1. The Kis-Balaton wetland including the site of the study, the Lake Fenéki (Fenéki Outer Reservoir and Grove Ingói).

When KBWPS was created, quantitative recording of its water budget has also begun with the traditional water balance equation, which includes the basic characteristics of water movements (*Zsuffa,* 1996). Members of the equation were determined from 1986 for Lake Hídvégi. In the case of Lake Fenéki, first calculations were made only after constructions have been completed, measuring points were installed, and the northwest part was flooded. Monthly water balance calculations for Grove Ingói and Lake Fenéki Outer Reservoir had started since 1993 and 2003, respectively. On the input side, the natural inflow (H), the precipitation (C), and the amount of water pumped (SZ) were taken into account, while on the output side, the outlet (L) and ET were listed. The change in water resources (ΔK) was calculated from the increase or decrease of the water level using the volume curve edited by the West-transdanubian Water Inspectorate. Subsurface water movements (inflow, leakage) are usually not quantified due to its negligible extent.

Water balance for Kis-Balaton (in the case of the three parts: Lake Hídvégi, Grove Ingói, and Lake Fenéki Outer Resevoir) is calculated separately in the following form:

$$(C + H + SZ) - (ET + L) \pm \Delta K = 0. \qquad (1)$$

In the present study, our aim was to determine the evaporation of Lake Fenéki (Grove Ingói + Lake Fenéki Outer Reservoir). Methods for calculating

ET for Hungarian areas are presented in numerous studies (*Tölgyesi*, 1993; *Kontur et al.,* 1993). When choosing a method, it is an important aspect, that data should be available from the regional meteorological parameters. The investigated period is 1970–2012, 43 years altogether. From the beginning of this period, meteorological measurements in the area of Kis-Balaton were made only at the stations in Zalaegerszeg, Keszthely, and Nagykanizsa. From these observations, air temperature and humidity data were available from 1970.

Wind speed and global radiation were available from the latter time. The West-transdanubian Water Inspectorate established automatic hydrometeorological stations (Balatonmagyaród Fekete Island, Balatonmagyaród beach 4T, Balatonmagyaród Almás Island, Keszthely-Fenékpuszta 21T) from the end of the 1990's. Air temperature, relative humidity, wind speed, and solar radiation were measured there. The data of the evaporation pan ("A" pan) at Balatonmagyaród beach 4T station are continuous from 1998.

The used data were as follows:

- Zalaegerszeg, Nagykanizsa (1986–1995): monthly potential evapotranspiration (*PET*) and *ET* based on the calculations of OMSZ (Hungarian Meteorological Service) (see later);
- Keszthely: monthly *PET* and ET (1970–1996; 2002–2012) on the basis of the OMSZ dataset;
- Keszthely: daily average temperature, relative humidity (1970–2012);
- Balatonmagyaród beach 4T, Balatonmagyaród Almás Island, Keszthely-Fenékpuszta 21T: daily average temperature, relative humidity, average wind speed (1993–2012).

Evaporation values (*PET, ET*) calculated by the operational soil moisture model (*Dunay,* 1993) developed for agrometeorological purposes was applied in this study. Data were taken from related OMSZ publications.

The long record of PET was approached in two ways. The Antal-formula which was developed for potential evapotranspiration in Hungarian climate conditions (*Antal*, 1968) was approached with the meteorological data of Keszthely for the whole period:

$$PET = 0.9(E - e)^{0.7}(1 + \alpha t)^{4.8} n \quad [mm / month], \tag{2}$$

where E is the saturation vapor pressure [hPa], e is the vapor pressure [hPa], α=1/273, t is the average air temperature [°C], and n is the number of the days in the month.

Monthly evaporation values were summarized annually. In the other method (earlier referenced as operational soil moisture model), data were taken from OMSZ agrometeorological publications for 1970–1996 on monthly basis. Data of 2002-2012 were summed from decade calculations for the water sector. For supplying the missing monthly data, we used linear regression using the

monthly values of the Antal-formula. Correlation before the supplement was 0.972.

For calculating *ET* at Kis-Balaton instead of developing a new "Kis-Balaton formula", which fits the local conditions, experts suggested using the Antal-formula for the determination of evaporation loss, although it was developed for Lake Fertő. The reason was that reed cover of both lakes was extensive. For water balance calculations, monthly *ET* was determined using the Lake Fertő formula for Grove Ingói as well:

$$ET = 0.42(E-e)^{0.9}(1+\alpha t)^{9}(1+0.015\cdot u)^{2}n \quad [mm/month], \tag{3}$$

where u is the monthly average wind speed [m/s]. To create this equation, energy balance measurements were made.

Evaporation increasing effect of wetland vegetation was taken into account with correction factors (crop coefficient) from April to October according to *Table 1*.

Table 1. Monthly average of crop coefficients for Kis-Balaton wetland

	Apr	May	Jun	Jul	Aug	Sep	Oct
Crop coefficients	1.02	1.11	1.2	1.26	1.21	1.13	1.11

For determination of local evaporation at Lake Fenéki Outer Reservoir, the method developed by *Dunay et al.* (1968) was used, which is based on "A" pan measurements:

$$ET = (100-H_{\%})(200-H_{\%})^{-1}\cdot t\cdot n \quad [mm/month], \tag{4}$$

where $H_{\%}$ is the monthly mean of relative humidity (%).

In hydrological practice, the water balance equation equals to 0, as the sum of positive and negative amounts. The expert calculating the equation needs a lot of data as many times as estimations are required. Flow rate of bigger streams is measured, some of them in every 15 minutes, but weekly or monthly measures are more common. At the most segments, water level series is created measuring the water level frequently, which is transformed into runoff time series. Integrating this series in time allows calculating the amount of inflow. Smaller streams are estimated based on hydrological analogy. At the outflow of Lake Fenéki (into Lake Balaton), some problems can temporarily occur because off

the strong upwind. In this case, an inverse circulation shows up at the surface and a kind of pulse is registered while measuring runoff. This phenomenon can be caused by the seiche of Lake Balaton. The quantification and determination of its effect on the amount of runoff should be solved in the future. In the case of precipitation over the lake, data of some precipitation gauges are considered. The precipitation of the 75 km^2 surface is calculated from the data of 3–6×200 cm^2 surfaces. Neither the surface nor the volume curves are renewed, although at some parts of Lake Hídvégi, sedimentation has begun. Curves of Lake Fenéki are sensitive; they change significantly in the error limit of water level registration. Accurate inventory of evaporation loss is not resolved, estimations containing error use of crop coefficients makes the review of vegetation rough. All of these can lead to significant errors. Experience has shown that water balance calculations are more difficult in the case of lower water level.

In a time series, within a normal data distribution, a two-tailed t-test was applied. To compare differences, when the Shapiro-Wilk normality test indicated a non-normal distribution, a non-parametric statistical hypothesis test, the Wilcoxon signed-rank test was used (SPSS Statistics v. 17.0; IBM Corp., New York, USA). To compare between daily *PET*s calculated on locally measured data observations at Keszthely, a linear regression ($y = a + bx$) was carried out, in which local *PET* was used as the dependent variable y and *PET* of Keszthely was the independent variable x. The fitness of curve was acceptable when the slope forced through the origin of the regression was close to 1 (*Alexandris* and *Kerkides*, 2003).

The purpose of the study was to develop ET prediction for Lake Fenéki. Originally the Box-Jenkins models were used for time-series analysis (*Box* and *Jenkins*, 1976) in the form of a seasonal auto-regressive integrated moving average (ARIMA(p,d,q)$(P,D,Q)_s$) model.

Box and *Jenkins* (1976) recommended the following general model:

$$\phi_p(B)\Phi_P(B)(1-B)^d\left(1-B^s\right)^D X_t = \theta_q(B)\Theta_Q\left(B^s\right)\alpha_t, \tag{5}$$

where d is the order of differencing, s is the length of the season, and D is the order of seasonal differencing.

The operator polynomials are as follows:

$$\phi_p(B) = \left(1-\phi_1 B - \ldots - \phi_p B^p\right), \tag{6}$$

$$\theta_q(B) = \left(1-\theta_1 B - \ldots - \theta_q B^q\right), \tag{7}$$

$$\Phi_p\left(B^s\right) = \left(1-\Phi_1 B^s - \ldots - \Phi_p B^{sp}\right), \tag{8}$$

$$\Theta_Q\left(B^s\right) = \left(1-\Theta_1 B^s - \ldots - \Theta_Q B^{sQ}\right), \tag{9}$$

$$(1 - B^s)X_t = X_t - X_{t-s}. \quad (10)$$

Among others, *Meshram et al.* (2011) applied seasonal ARIMA model to project evaporation for India (Solapur Station). The original *ET* time series for Lake Fenéki has not demonstrated any trend; data exhibited numerous peaks, which appeared to be equally spaced. This phenomenon suggested the presence of a periodic component to the time series, where the peak took place in the growing seasons.

To determine ARIMA model for time series analysis and produce forecasts, the SPSS 17.0 statistical program was applied. The program includes an expert modeler modul that automatically estimates the best-fitting ARIMA model, eliminating the need to identify an appropriate model through trial and error.

There are three basic terms of ARIMA models; auto-regression (AR), differencing or integration (I) and moving-average (MA) components. Although these components respond differently to a random disturbance, they are based on the concept that they may be described by ARIMA models. Since time series showed seasonality, the seasonal ARIMA $(p,d,q)(P,D,Q)_s$ orders were presented. An effective approach for isolating seasonal orders is to calculate the autocorrelation functions (ACF) and partial autocorrelation function (PACF) plots at the seasonal lags (http://www.spss.com). The number of spikes showed the order of auto-regression.

3. Results and discussion

The Kis-Balaton lake system, in which Lake Hídvégi is about 18 km^2, is typically an open water area (85%). The cover of marsh vegetation is about 15%. The area of Grove Ingói is about 16 km^2. The distribution of open water, reed, bulrush, and sedge shows a mosaic structure. Here, the share of open water is only about 15%. The Lake Fenéki Outer Reservoir – the non-flooded area – is about 41 km^2. The permanent water cover is low, about 1–2%, but temporarily it can be higher (even almost the whole area), belonging to the shallow water covered, wetland types.

A ten-year period was chosen to overview the evaporation conditions of the wider environment, which includes our study site as well. The calculated monthly values of *PET* and *ET* between 1986 and 1995 at the three sites near Lake Balaton (Keszthely, Zalaegerszeg, Nagykanizsa) matched only in respect of *ET*. *PET* values of the three stations differed significantly ($p \leq 0.0001$). The annual sums ranged between 840 and 1120 mm. The average of the ten years is the highest in Kesztely (1060 mm) as expected. This is followed by Zalaegerszeg (1025 mm) and Nagykanizsa (940 mm). Monthly maximum of *PET* occurred in August 1992 (245 mm) in Keszthely. The

maximum of *ET* (144 mm) calculated by the OMSZ operational soil moisture model occurred in July 1991, the annual sums are between 460 and 760 mm, but the differences are not significant ($p \leq 0.064$–0.502), while the average of the 10 years is about 630 mm from the values of the three stations. Even when the *ET* values from OMSZ operational soil moisture model are accepted for the wider environment, they are not applicable for Kis-Balaton because off its moist endowments. *ET* values of wider environment are surmised to be underestimated.

Data from Keszthely was included in the *PET* studies for the longer period (1970–2012), since this is the closest station to the study area (Kis-Balaton) and it is situated at the side of the prevailing wind direction.

In the first step, *PET* values for Kis-Balaton between 1970 and 2012 were determined by the Antal-method, and they were compared to the results of OMSZ operational soil moisture model. *PET* of Kis-Balaton was determinable between 1993 and 2012 by Antal-formula with locally measured data. Correlation coefficient of calculated annual *PET* between using the Antal-formula and the soil moisture model of the OMSZ was acceptable, 0.781. It should be noted that in this period in the last, shorter section (1994–2012), the correlation coefficient was much higher (0.936). The stronger correlation can be explained by changes in observation methodology (from the early 90s, the manual observation was replaced by automatic weather stations). Looking at the annual *PET* data series (*Fig. 2*) it can be stated, that the nature of changes is similar, differences among the three curves are small, but sometimes significant. The annual *PET* calculated from the data of Keszthely is 5.2% ($p \leq 0.012$) higher than the *PET* calculated from the data of Kis-Balaton. Similarly, PET calculated from the OMSZ soil moisture model is 4.1% ($p \leq 0.008$) lower, than *PET* calculated from the data of Keszthely. There was no significant difference between *PET* calculated from the local data of Kis-Balaton and the data of the OMSZ model ($p \leq 0.54$). The difference between the annual *PET* sums from the two procedures is within ±140 mm. Values calculated from the data of Kis-Balaton headed between the two other curves, close to them. Years between 2000 and 2003 should be highlighted. At this period (very hot and dry), the values calculated from the local data are much lower. *PET* values based on local data were the lowest of all; the Antal-formula and the OMSZ operational soil moisture model exceeded *PET* with 110–261 mm and 73–167 mm, respectively. The reason might be the higher humidity values in the area of Kis-Balaton (wetland).

The slope of linear trend fitted to the data of the 43 years calculated for the data of Keszthely showed 3.7 mm increment per year ($p \leq 0.004$), while according to the OMSZ model it is 2.8 mm increase per year ($p \leq 0.046$). The linear trend forecast fitted to the data of Kis-Balaton was not significant ($p \leq 0.197$).

The most important statistical indicators of *PET* data set calculated by different approaches are summed in *Table 2*.

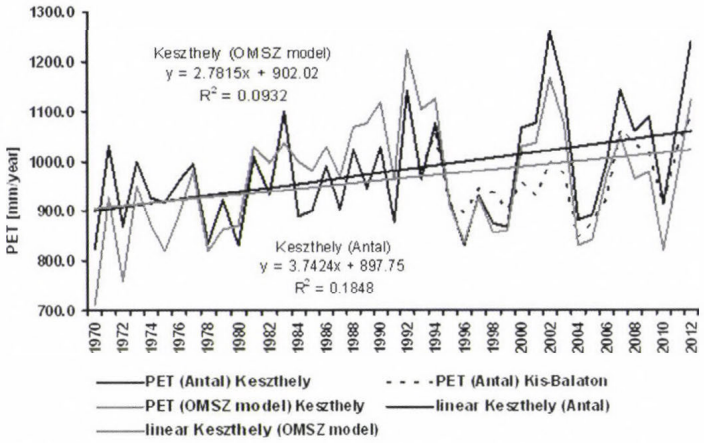

Fig. 2. Comparison of potential evapotranspiration (*PET*) calculated by using data of (i) Keszthely [*PET* (Antal) Keszthely]; (ii) the soil moisture model of the OMSZ (Hungarian Meteorological Service), and (iii) locally measured meteorological data [*PET*(Antal) Kis-Balaton]. We expressed PET after empirical equation processed by *Antal* (1968).

Table 2. Statistical parameters of potential evapotranspiration (*PET*)

	Avg.	Min.	Max.	SD	CV%
PET (Antal) Keszthely (1970–2012)	980	824	1261	109	11%
PET (Antal) Kis-Balaton (1993–2012)	963	844	1089	68	7%
PET (OMSZ model) Keszthely (1970–2012)	963	711	1224	114	12%

The monthly variation of *PET* within the year was analyzed, too (*Fig. 3*).

The maximum monthly *PET* was expected in July, the minimum in December. The difference between the two calculation methods was 7–10 mm in March, September, October, and under 6 mm in the other months. The share of the season from the annual PET was 740 mm, 75%.

Correlation among *PET* values calculated from the data of Keszthely and Kis-Balaton was very strong between 1993 and 2012, the correlation coefficient was 0.998. Thus the data of Keszthely can be converted into the data of Kis-Balaton (*Fig. 4*). There was no significant *SE* when the monthly evaporation sum was below 50 mm, however when it exceeded 100–150 mm, the *SE* was much higher weakening the accuracy of *ET* estimation.

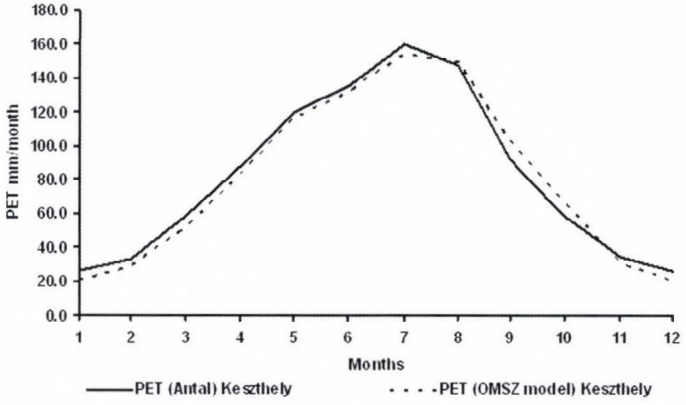

Fig. 3. Monthly means of potential evapotranspiration of Keszthely using the Antal-formula and the soil moisture model of the OMSZ between 1993 and 2012.

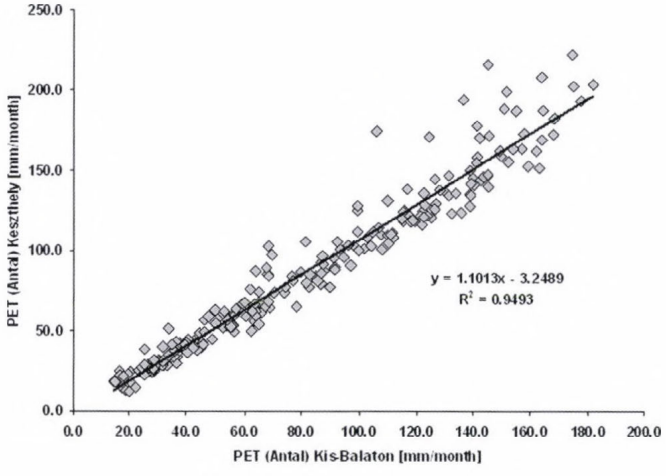

Fig. 4. Relationship between monthly calculated evapotranspiration from Keszthely meteorological data and locally measured meteorological data. To calculate *PET*, the Antal-formula was applied. The number of observed pairs was 240.

The estimation of *ET* for the area is a much more complex task, than the *PET* projection. Evaporation loss of water covered parts in the evaporation of Lake Fenéki is considered close to the potential, while at non-flooded parts, *ET* data seems to be useable. The *ET* of Lake Fenéki calculated by the OMSZ soil moisture model was not acceptable on annual basis (this model has been

developed for average soil conditions, not for quasi wet environment). Regional evaporation of Lake Fenéki was examined between 1970 and 2012. The whole area is consistent for the years 1970–1992 in determining the evaporative water loss, since this is a period before the construction (partial flood). Since 1993 two parts of the lake are distinguished: (a) the flooded Grove Ingói and (b) the non-flooded (Outer Reservoir) Lake Fenéki.

The length of *ET* time series of Grove Ingói controlled by water balance was a 20-year period (*Fig. 5*). The values of annual *ET* were between 730 and 1070 mm, with an average of 890 mm, which was 70 mm, 8% ($p \leq 0.0001$) less, than those of *PET*.

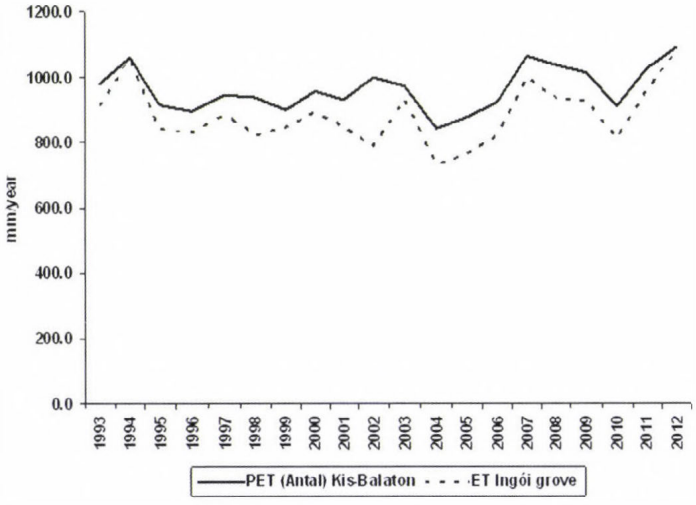

Fig. 5. The actual (*ET*) and potential (*PET*) evapotranspiration of Grove Ingói.

The estimation of evaporation for Lake Fenéki Outer Reservoir was prepared as well. Here Lake Fertő formula can not be used, since there is no permanent water cover. Water supply varying in time and space is typical for this area, which can not be followed exactly from the measured atmospheric data. Aerial photos were taken at the area (1988, 1992, 1993, 1994, 1995, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2008). According to them, the classification of the vegetation distribution (*Table 3*) was made as a starting base (*Pomogyi*, 2001; *Pomogyi et al.,* 1996). The year of 1992 is considered to be typical for the period before taking the aerial photos. 1992 was chosen in analysis, because the photos of 1988 were not detailed enough and the mapping methods were not up-to-date.

Table 3. Vegetation map data of Lake Fenéki Outer Reservoir

Lake Fenéki Outer Reservoir	Avg [ha]	%	min	max	SD	CV%
Reed	1 238	30%	1 117	1 363	74	0
Other wetland crops (marsh macrophytes)	1 753	43%	1 459	2 075	196	11
Swamp forest	177	4%	143	222	20	0
Herbaceous crops and grasses	387	9%	324	553	57	15
Wood	471	12%	309	588	114	24
Open water with seaweed	63	2%	12	99	27	42
Altogether	4 090	100%	4 064	4 107	12	0

According to the values of the table, open water surface occupies small area (2%). About 80% of the vegetation is considered to be well stocked with water similarly to wetland habitats. (In the especially dry years, lower percentage is expectable). For the characterization of the period between 1970 and 2012, precipitation data from Keszthely were used. The years were defined to be dry, when the lack of precipitation of the vegetation period exceeded the average of the period between 1970 and 2012 with 20%. These years were also warmer, than the average; their annual mean average temperature was 0.4 °C higher than the long-term average. Nine dry-warm years were found (1971, 1977, 1981, 1984, 1988, 1993, 2000, 2001, 2011), which is the fifth part of the whole period.

In the present processing, Eq. (4) was used for the estimation of *ET*. *Dunay et al.* (1968) based this method on the measurement data of "A" class evaporation pan, and suggested for the calculation of *PET*. This method was chosen, because the most of the area is a wetland well supplied with water, so a well-approximated estimation is expected. (Please note that *ET* was estimated by *Varga* (2005) for the non-flooded area, using *PET* values of the Antal-formula and the formula of *Turc* (1961) with a factor of 0.75 and 0.25.)

Application of the formula (4) was a good choice. The accuracy of calculated monthly *ET* values were checked using the monthly water balances for the period of 2003–2012. (Meteorological data used for the calculations were taken from the local measurements.) Estimation due to *PET* was quasi negated with the result that values measured with "A" class pans lag behind potential estimations. The average of the studied 10 years was 703 mm.

The evaporation of Grove Ingói and Lake Fenéki Outer Reservoir (real values, controlled with the water balance) was compared with the values of the "A" class pan at the study site (*Fig. 6*). *ET* of Grove Ingói between April and October (1998–2012) exceeded the *ET* calculated for Lake Fenéki Outer Reservoir with 14.4% ($p \leq 0.0001$). The data of "A" pan and *ET* calculated for

Grove Ingói differed with 24.1% ($p \leq 0.0001$) in the vegetation period between 1998 and 2012. The data of "A" pan and *ET* calculated for Lake Fenéki Outer Reservoir differed less, with 7.9% ($p \leq 0.0001$), as expected. *ET* is significantly underestimated in summer and autumn in the "A" pan, so its usage is not suggested for the estimation of evaporation at aquatic habitats.

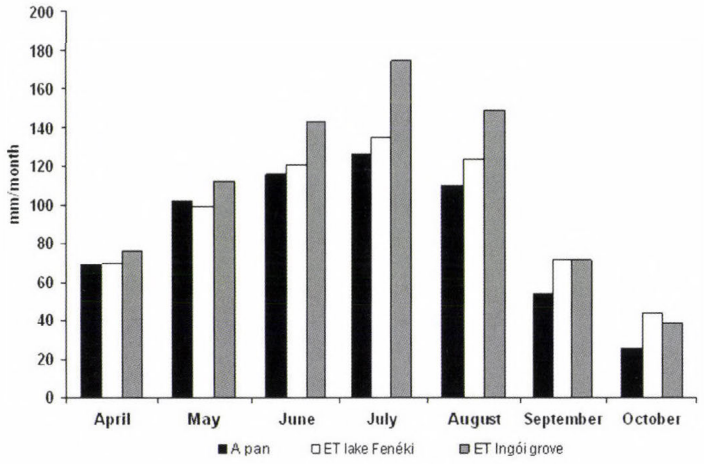

Fig. 6. Water losses calculated from water budget for Grove Ingói, Lake Fenéki Outer Reservoir and locally observed „A" pan evaporation (1998–2012).

Hereinafter, our calculations were extended to the period before the flood. The accuracy of *ET* calculated with Eq. (4) for Lake Fenéki Outer Reservoir was checked with the water balance, so this equation was adopted to the earlier period for the non-flooded area as well. Between 1970 and 2002, *ET* was calculated on a monthly basis. (The measured data of the OMSZ station in Keszthely were used.) Dataset checked with water balance was added to the values calculated with Eq. (4) for Grove Ingói before the flood (1970–1992). So a monthly data set for the period between 1970 and 2012 was made, from which annual *ET* was calculated. At Lake Fenéki Outer Reservoir for the whole period (1970–2012), *ET* was calculated with the Eq. (4). The data of the period between 1970 and 2012 were checked with the water balance. *ET* data set for the whole Lake Fenéki was made as well. Actual evaporation of Lake Fenéki for the period between 1970 and 2012 was derived from the calculated *ET* for Grove Ingói and Lake Fenéki Outer Reservoir, where the area ratio was taken into account (*Fig. 7*).

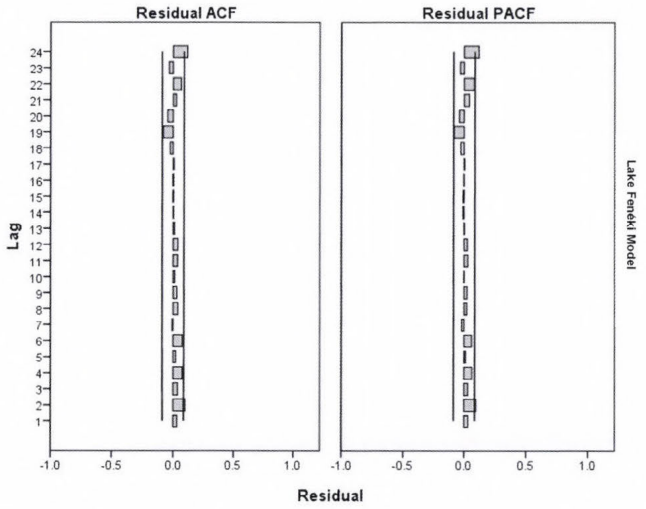

Fig. 7. Residuals (ACF and PACF) from the fitted ARIMA $(0,0,3)(0,1,1)_{12}$ for ET of Lake Fenéki.

The average annual evaporation of the whole period is 809.3±87.8 mm. If the nine dry-warm years are highlighted, their average annual evaporation exceeds the long-term average (874.7±37.6 mm) with 78 mm, with significantly lower standard deviation values. The average evaporation of the remaining wet-cold years is 796.6±89.4 mm. These years dominated (36 years), so sum of evaporation and standard deviation were close to the long-term average. None of the linear trends fitted to the two different periods (before and after the flood) were significant (1970–1992: $p \leq 0.144$; 1993–2013: $p \leq 0.366$).

In order to study *ET* time series of Lake Fenéki, an ARIMA model has been processed. We assumed that the process of *ET* was stochastic in nature. Therefore, we attempted to investigate the applicability of autoregressive integrated moving averages (ARIMA) modeling, a special time-series technique for developing forecast model for *ET*. In our study, the partial autocorrelation function (PACF) produced peak at lag 12, meaning that seasonal ARIMA $(0,0,3)(0,1,1)_{12}$ model is suggested to be the best fit for *ET* time series analysis.

To summarize *ET* seasonal variations in the present investigation, the SPSS time series modeller (http://www.spss.com) has shown that the seasonal ARIMA $(0,0,3)(0,1,1)_{12}$ model including additional seasonal term simply multiplied with the non-seasonal term has to be chosen. The reason of our choice was due to the fact that this model produced the lowest RMSE (quadratic scoring rule that measures the average magnitude of the error) and MAPE (mean absolute percentage error), equaling to 14.273 and 21.1%, respectively. The Ljung-Box-statistic (*Ljung* and *Box*, 1978) provided insignificant value of 0.398 meaning that there is no structure in the observed series, which is not accounted for by the

model. For a model to be considered as adequate at describing evaporation time series, the residuals of the model have to be correlated; both ACF and PACF have to lie within limits counted by equation (*Fig. 7*). The residuals should be without pattern. As the Kolmogorov-Smirnov-test of noise residual indicated normal distribution, the multiplicative decomposition of time series was successful ($p \leq 0.535$). The estimated model parameters for ARIMA $(0,0,3)(0,1,1)_{12}$ computed by SPSS expert modeler are presented in *Table 4*.

Table 4. Estimated model parameters for seasonal ARIMA $(0,0,3)(0,1,1)_{12}$

				Estimate	SE	t	Sig.
Lake Fenéki Model	Square Root	MA	Lag 1	–0.224	0.043	–05.161	0.000
			Lag 3	–0.177	0.043	–4.065	0.000
		Seasonal Difference		1			
		MA, Seasonal	Lag 1	0.952	0.028	33.913	0.000

Based on RMSE value, seasonal ARIMA $(0,0,3)(0,1,1)_{12}$ was selected for forecasting of monthly *ET*, at Lake Fenéki (*Fig. 8*). *Fig. 8* also includes the upper and lower (UCL, LCL) values of an estimated confidence interval at 95% for the predictions. The chosen projected time period was only three years (2013–2015).

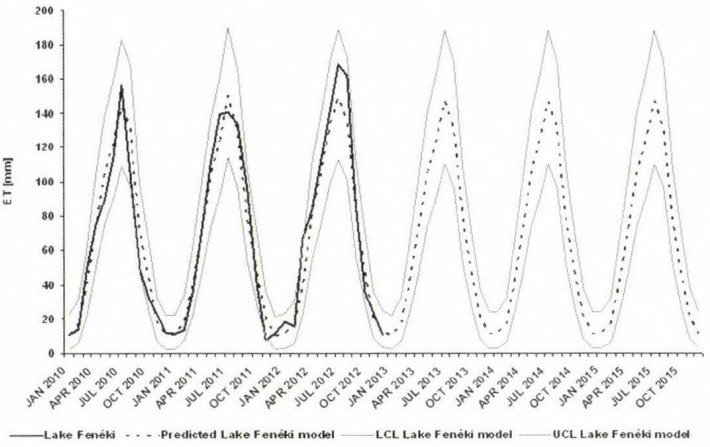

Fig. 8. The past and the future of the *ET*. Pattern of three-year actual and predicted *ET* data with three-year forecast (2013-2015) for monthly *ET* sums of Lake Fenéki by using the ARIMA $(0,0,3)(0,1,1)_{12}$ model. The LCL and UCL are the lower and upper limits of the confidence interval (95%), respectively.

The ARIMA $(0,0,3)(0,1,1)_{12}$ model parameters finalized for *ET* forecasting were as follows: $\theta_1 = -0.224$, $\theta_3 = -0.177$, $\Theta_1 = 0.952$. In our case $D=1$ (the order of seasonal differencing), $s=12$ (number of month per year). Seasonal pattern of *ET* series was also maintained in the projected values (see also *Table 4*).

4. Conclusions

Investigations were carried out on evaporation of Lake Fenéki (Kis-Balaton wetland) in the time period between 1970 and 2012. *PET* for the lake was generated using meteorological data of different stations. This study revealed that the locally measured meteorological values or regression equation derived from meteorological data of Keszthely would be appropriate in Lake Fenéki's *PET* calculation.

One part of the lake's observation site (Grove Ingói) has been artificially flooded in 1993. Irrespective to human intervention, the whole lake's spatial *ET* was statistically the same during the two time periods with modified watering levels (unflooded and flooded Grove Ingói). The long-term yearly *ET* averages of Lake Fenéki were 813 and 805 mm/year between 1970–1992 and 1993–2012, respectively. On the basis of fitted curves (*ET*), no significant trend has been confirmed either for the whole time period or after/before artificial intervention. Using trend analysis, we might conclude that in spite of flooding of Grove Ingói, there was no significant difference in long-term yearly *ET* sum of Lake Fenéki between 1972 and 2012. The probable reason might be the shortage of the used models. These formulas could not distinguish the various kinds of crop covers and their different evapotranspiration.

ET values of Lake Fenéki showed seasonal cycle, therefore, monthly data were used for generating a stochastic model. Investigation indicated that the seasonal ARIMA $(0,0,3)(0,1,1)_{12}$ model is a viable tool for studying long-term monthly *ET* data for Lake Fenéki. The generated *ET* values kept the earlier observed seasonal pattern, probably due to special wetland circumstances. The seasonal ARIMA $(0,0,3)(0,1,1)_{12}$ with lowest RMSE may also be selected for forecasting of monthly *ET* sums at Lake Fenéki (part of Kis-Balaton wetland, Hungary) using the following equation:

$$\left(1-B^{12}\right)X_t = \left(1-0.224B\right)\left(1-0.177B\right)\left(1+0.952B^{12}\right)\alpha_t, \qquad (11)$$

where X_t represents the time series data at period t. α_t represents a Gaussian white noise process (random shock) at period *t*. *B* represents a backward shift operator (*Box* and *Jenkins*, 1976).

Future research should be addressed to extend our results on other regions of Kis-Balaton wetland.

Acknowledgements: Present article was published in the frame of the project TÁMOP-4.2.2.A-11/1/KONV-2012-0064. The project is realized with the support of the European Union, with the co-funding of the European Social Fund.

References

Alexandris, S., Kerkides P., 2003: New empirical formula for hourly estimations of reference evapotranspiration. *Agr. Water Manage. 60,* 157–180.

Antal, E., 1968: Új módszer a potenciális evapotranspiráció számítására. *Beszámolók 1968*, Országos Meteorológiai Szolgálat Hivatalos Kiadványai, Budapest. (In Hungarian)

Antal, E., Baranyi, S., and *Kozmáné Tóth, E.*, 1977: A Balaton hőháztartása és párolgása. *Hidrológiai Közlöny 57*, 182–190. (In Hungarian)

Box, G.E.P. and *Jenkins, G.M.*, 1976: Time series analysis, forecasting and control. Holden Day, San Francisco.

Box, G.E.P., Jenkis, G.M., and *Reinsel, G.C.*, 2008: Time series analysis, Forecasting and Control. 4th edn. John Wiley and Sons, Inc, New Jersey.

Box, G.E.P., and *Jenkins, G.M.*, 1976: Time series analysis, forecasting and control. Holden Day, San Francisco.

Brockwell, P.J. and *Davis, R.A.*, 2001: Introduction to Time Series and Forecasting. Springer Verlag, New York.

Dunay, S., 1993: A termőhely vízháztartásának modellezése. In (Ed.: Szalay) Éghajlati és agrometeorológiai tanulmányok 1) Országos Meteorológiai Szolgálat Hivatalos Kiadványai, Budapest. (In Hungarian)

Dunay, S., Posza, I., and *Varga Haszonits, Z.*, 1968: Egyszerű módszer a tényleges evapotranspiráció és a talaj vízkészletének meghatározására. I. A párolgás meteorológiája. *Öntözéses gazdálkodás 6*, 39–48. (In Hungarian)

Harkay, M. 1983: A Kis-Balaton rekonstrukció és környezeti hatásai. Évezredek üzenete a láp világából. Régészeti kutatások a Kis-Balaton területén, Kaposvár-Zalaegerszeg, 7–10.(In Hungarian)

Hatvani, I.G., Kovács, J., Székely Kovács, I., Jakusch, P., Korponai, J. 2011: Analysis of long-term water quality changes in the Kis-Balaton Water Protection System with time series-, cluster analysis and Wilks'lambda distribution. *Ecol.Eng.* 37, 629–635.

Havalda, E., 1930: A Balaton párolgása *Vízügyi Közlemények 1,* 87–97. (In Hungarian)

Istványovics, V., Kovács, A., Vörös, L., Heródek, S., and *Pomogyi, P.*, 1997: Phosphoruscycling in a large, reconstructed wetland, the lower Kis-Balaton Reservoir (Hungary). *Int.Ver.Theor. Ang.Limn.26,* 323–329.

Ljung, G.M. and Box, G.E.P., 1978: On a Measure of a Lack of Fit in Time Series Models. *Biometrika 65*, 2, 297–303.

Kontur, I., Koris, K., and *Winter, J.*, 1993: Hidrológiai számítások. Akadémiai Kiadó, Budapest. (In Hungarian)

Korponai, J., Braun, M., Buczkó, K., Gyulai, I., Forró, L., Nedli, J., and *Papp, I.*, 2010: Transition from shallow lake to a wetland: a multi-proxy case study in Zalavari Pond, Lake Balaton, Hungary. *Hydrobiologia 641*, 225–244.

Kovács, J., Hatvani, I.G., Korponai, J., and *Kovácsné, Sz.I.*, 2010: Morlet wavelet and autocorrelation analysis of long term data series of the Kis-Balaton Water Protection System (KBWPS). *Ecol. Eng. 36,* 1469–1477.

Meshram, D.T., Gorantiwar, S.D., and *Lohakare, A.S.*, 2011: ARIMA model for forecasting of evaporation of Solapur station of Maharashtra, India. *Mausham 63*, 573–580.

Nguyen, H.L., Leermakers M., Kurunczi, S., Bozó, L., Baeyens, W., 2005: Mercury distribution and speciation in Lake Balaton, Hungary. *Sci. Total Environ. 340,* 231–246.

Nováky, B., 2005: A Balaton vízpótlása és éghajlata. *Vízügyi Közlemények 87,* 1, 114–123.(In Hungarian)

*Padisak, J.*and *Reynolds, C.S.*, 2003: Shallow lakes: the absolute, the relative, the functional and the pragmatic. *Hydrobiologia 506–509*, 1–11.

Pomogyi, P., 2001: Az Ingói berek vegetáció térképezése 1998 és 2000 között. Nyugat-dunántúli Vízügyi Igazgatóság: Szombathely. (In Hungarian)

Pomogyi, P., Szeglet, P., and *Csato, E.*, 1996: The change in the reed community of the KBWPS, Feneki pond, based on vegetation-mapping results. In (Ed.: *Pomogyi, P.*) 2. Kis-Balaton Conference: Summary of the Research Result on the KBWPS between 1991–1995. University of Pannonia Georgikon, Faculty of Agriculture, Keszthely, 206–207.

Psilovikos, A. and *Elhag, M.*, 2013: Forecasting of Remotely Sensed Daily Evapotranspiration Data Over Nile Delta Region, Egypt. *Water Res. Manage 27,* 4115–4130.

Ramsar Convention, 1971: *Convention on Wetlands of International Importance especially as Waterfowl Habitat.* Ramsar (Iran), February 2, 1971. UN Treaty Series No. 14583. As amended by the Paris Protocol, December 3, 1982, and Regina Amendments, May 28, 1987.

Szesztay, K., 1962: A Balaton vízháztartása. *VITUKI Tanulmányok és eredmények 9*, 299–310. (In Hungarian)

Tátrai, I. Mátyás K., Korponai J., Paulovits G., and *Pomogyi P.*, 2000: The role of the Kis-Balaton Water Protection System in the control of water quality of Lake Balaton. *Ecol. Eng. 16,* 73–78.

Tölgyesi, L., 1993: Az éghajlat változékonyságának hatása a talajnedvességre. In (ed.:*Szalay S.*) Éghajlati és agrometeorológiai tanulmányok 1. Országos Meteorológiai Szolgálat Hivatalos Kiadványai, Budapest. (In Hungarian)

Turc, L.,1961: Evaluation des besoins en eau d'irrigation, evapotranspiration potentielle, formuleclimatique simplifee et mise a jour.*Annal Agronomy 12*, 1, 13–49. (In French)

Varga, Gy., 2005: A Balaton vízháztartási viszonyainak vizsgálata. *Vízügyi Közlemények 87*, 100–101. (In Hungarian)

Varga, Gy., Pappné, U.J., and *Bálint, G.*, 2007: Recent extremes in the water budget of Lake Balaton. *Georgikon for Agric. 10,* 1, 25-44.

Virág, Á., 1997: A Balaton múltja és jelene. Eger Nyomda Kft., Eger. (In Hungarian)

Zsuffa, I., 1996: Műszaki hidrológia. Műegyetemi Kiadó, Budapest. (In Hungarian)

URL addresses:

http://www.spss.com

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 1, January – March, 2015, pp. 111–126

Evaluation of the cold drops based on ERA-Interim reanalysis and ECMWF ensemble model forecasts over Europe

Nikolett Gaál[1] and **István Ihász**[2]

[1]*Department of Meteorology, Eötvös Loránd University*
P. O. Box 32, H-1518 Budapest, Hungary
E-mail: gaalnikki@gmail.com

[2]*Hungarian Meteorological Service,*
P.O. Box 38, H-1525 Budapest, Hungary
E-mail: ihasz.i@met.hu

(Manuscript received in final form February 19, 2014)

Abstract—In our work, we planned deeper understanding of cold drops, closed air masses separated from the main western stream, by using ECMWF ERA-Interim reanalysis and ensemble forecasts. Upper level low (ULL) recognition algorithms were used to study 70 independent cold drop occurrences from the last decade in the middle and eastern European region. This led to the ascertainment of the usual location of ULLs in relation to Hungary, their core temperature, axis lean, horizontal temperature change, and identification on „plum" diagrams. Our studies included the usage of the potential temperature of the 2PVU (2 potential vorticity unit) surface, potential vorticity field related to 315 K, and 300 hPa wind speed These recommended new variables are available from operational deterministic and ensemble forecasts, and their usage is highly effective, hence making the identification of cold drops a lot easier than before.

Key-words: cold drops, upper level lows, reanalyses, ensemble model, statistical studies, 2PVU, isentropic potential vorticity, jet stream, visualization, case studies

1. Introduction

Synoptical studies of the cold drops got less attention in the past decades (both in domestic and international literature) than their significance would suggest. Therefore, we have a slice of all the information about their formation points, development conditions, synoptical and dynamical backgrounds. We find that the importance of their study lies in their ability to inflict natural disasters, such as massive thunderstorms that can lead to flooding, or in rare cases, even

tornadoes. Furthermore, upper level lows, or ULLs also have a negative impact on human health. At first, the expression of cold drop, ("der Kaltufttropfen") was likely used by *Scherhag* (1948). His article contained case studies, including winter and summer weather situations influenced by cold drops above the territory of Germany.

A cold crop (also called upper level low or ULL) is a closed air mass separated from the main western stream. Isolated from the cooler air of higher latitudes, it carries air substantially colder than its surroundings, to the warm regions of lower latitudes. In practice, its analysis is done at heights of about 5500 m, and a pressure level of 500 hPa, in the middle troposphere. ULLs being elliptically shaped with a diameter of hundreds kilometres, they resemble miniature cyclones on satellite images. A cold drop can determine the weather of the region for a couple of days, often bringing high amounts of intensive rain, especially in summer. They can occur at any time of the year. The unstable nature of ULLs provides perfect conditions for the forming of hails and thunderstorms in summer, and for intensive snowing winter. One of the main characteristics of a cold crop is that the air inside its higher regions is a lot colder than that of the area outside the cold drop. ULLs are sometimes referred to as "eddies", because of their counter-clockwise cyclone-like motion.

Peltonen (1963) investigated an intensive upper air low occurred in late autumn above Northern Europe. He found that the direction of the transition was parallel to surface wind direction, but the typical transition wind speed was only about 60–70% of the surface wind speed.

In Hungary, the first comprehensive study was made by *Bodalainé* (1983). She created subjective classification system based on synoptical patterns caused heavy floods in the Carpathian Basin.

A quite complex study was made by *Kurz* (1990) for the territory of Germany. Kurz investigated the life cycle of the cold drops and studied the relationship between the cold drops and cut off cyclones. He found that this phenomenon mainly occurred in winter the season. He stated that two main mechanisms could cause cold drops. The first one is the consequence of the cut off procedure, and the second one is the consequence of the cyclogenesis in the case when, due to dissipation near to the surface upper level, low becomes dominant.

It is important to note that cold drops could occur in Mediterranean areas too, and they are even mentioned in studies in central and northern Europe. Impacts of the cold drops are important in the Iberian Peninsula, too (*León*, 2003). In the Iberian Peninsula, especially in the southeastern part of Spain it occurs when a large size polar air mass slowly moves to that region in the middle and upper troposhere. It causes high impact weather including heavy rains or snowing (up to 500 mm/day) and extremely strong wind (between 100 and 200 km/h), when cold air mass mixes with warm Mediterranean air mass.

Complex climatological study based on a 41 years long NCEP reanalysis was made by *Nieto et al.* (2005). An objective recognition system using three meteorological variables was applied. The three meteorological variables were the following: cold core at 200 hPa, circulation pattern at 200 hPa, and thermal front parameter field. In contrast to the former studies, it was found that cold drops occurred more often in the winter season than in summer. Authors also found a connection between jet streams and upper level lows.

2. Life cycle of the ULL and relationship between blocking and upper level lows

A classic cold crop has four development phases, starting with the ULL isolating from the main stream and ending with dissolving or fusing with another stream (*Nieto et al.*, 2005). The four phases are

— ULL,
— tear-off,
— cut-off, and
— final stage.

2.1. Upper level low phase

In order for a cold drop to form, there have to be unstable waves inside the main stream, where the temperature wave is behind the geopotential wave (*Fig. 1a*). This is the phase where the ULL is still behind the frontal cloud mass, so it shows as a clearly visible cloud trail on satellite images.

2.2. Tear-off phase

The main meteorological process of this phase is the trough tearing off from the main stream (*Fig. 1b*). The waves amplitude gets higher ("the wave gets deeper"), followed by cold air detaching from the stream in its southern regions. The bottom part of the ULL slowly isolates from the main stream, leading to a closed circulation in the upper troposphere.

2.3. Cut-off phase

Contrary to the earlier phases, the isolation is completely finished in this phase, and the ULL is at its prime. The wind field shows the most advanced closed circulation at 500 hPa (*Fig. 1c*).

2.4. *Final phase*

Convection begins to develop in the cold drop (except for its coldest parts). As the air on the surface gets warmer, the eddies' circulations get slowed down by friction. The upper air mass gets warmer by the conviction and friction starts to take effect in this region as well. This causes the ULL to decay slowly (*Fig. 1d*). In most cases, the ULL reattaches to the main western stream before dissolving completely (*ZAMG*, 2007).

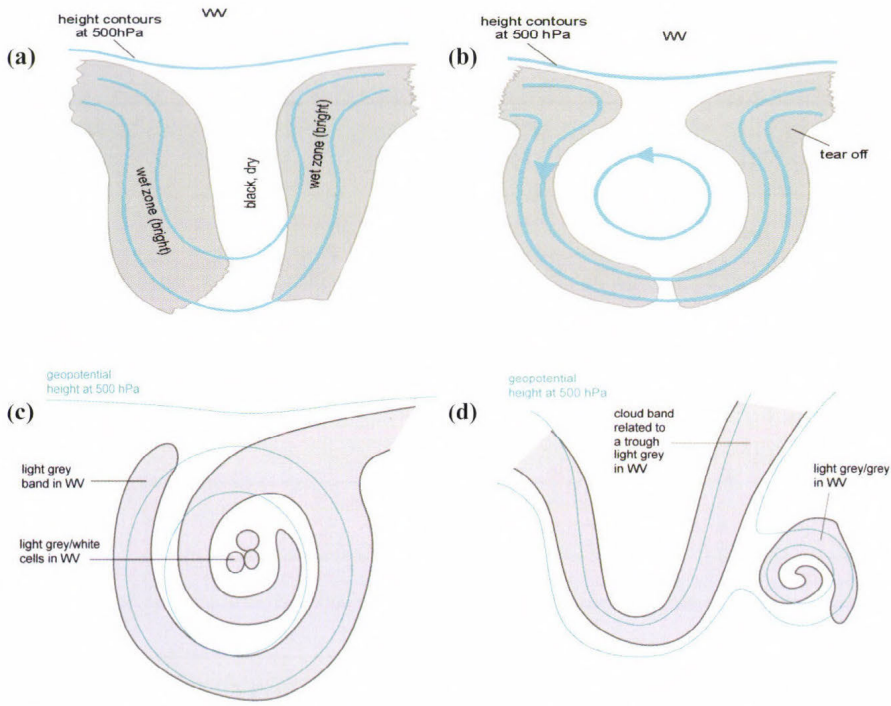

Fig. 1. Phases of the ULL: upper level low phase (a), tear-off phase (b), cut-off phase (c), final phase (d) (*ZAMG*, 2007)

The development of ULL usually takes 3–10 days. There are two kinds of ULLs, based on size and lifetime, "small eddies" which last 2–4 days, and "big eddies" which can last up to 14 days. Larger eddies are usually more common than smaller ones.

When studying cold drops, the phenomenon of blocking has a very high synoptic significance. First comprehensive studies of the blocking were made by *Rex* (1950a, 1950b). At the middle latitudes, two main patterns of the flow are typical: zonal flow and meridional flow (blocking). Life cycle of the blocking can extend from a few days up to two weeks. In the typing of the blocking

situations, we can see characteristic flow images that can be analyzed well at 500 hPa pressure level during the occurrence of blocking events. By analyzing the blocking anticyclones at 500 hPa, the following types can be distinguished: Rex-type blocking, Omega-situation, "fire ring" or detaching anticyclone, detaching low-pressure system, and splitting flow field. For our studies, Omega-type blocking is the one of important, because this event often leads to the detachment of cold drops (*Pelly* and *Hoskins*, 2003).

3. Former studies on cold drops in Hungary

ULLs are mentioned only in a few pieces of Hungarian literature. A synoptic classification system for the Carpathian Basin was invented by *Bodolainé* (1983). This classification system listed situations only that caused floods in the area of the rivers Danube and Tisza. She specified 7 classes: western, "western boundary disturbance", zonal, passing Mediterranean, central, cold air drop, and western cyclone-type. *Bodolainé* was a pioneer (both in Hungary and Europe in general) in studying the relation between dangerous weather conditions and cold air drops. She found that cold air drops occur quite rarely and mainly in the summer season.

Their occurrence often forms mezo-scale convective weather-systems (*Horváth*, 2007), and in some cases they have strong effect on the flooding of the river Tisza (*Bodolainé*, 1983; *Bonta et al.*, 1989; *Szépszó*, 2003).

The distinction between upper level lows and upper level cold vortexes is typical. Due to this distinction, the characteristic size of the upper cold drops is up to thousand kilometers, while the characteristic size of the cold vortexes varies between 1000 and 3000 kilometers. Due to the meteorological terminology phrase, "cold drop" is used if around cold core there are isolated isolines in geopotential field in the middle troposhere and the characteristic size is only several hundred kilometers (*Fig. 2*).

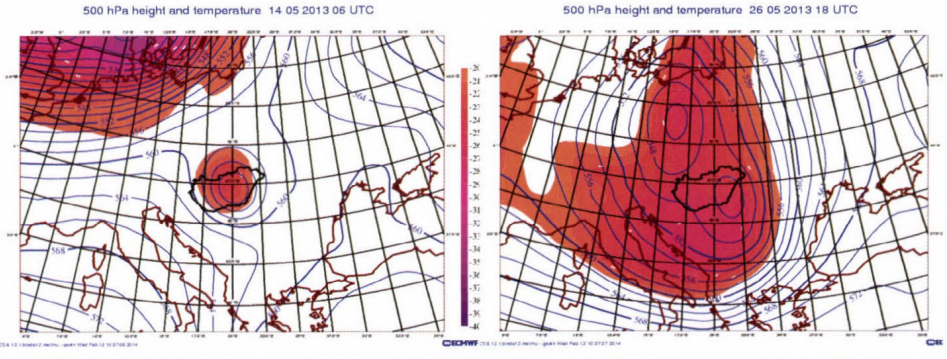

Fig. 2. Cold drop: May 14, 2013, 06 UTC (left), upper level cold vortex: May 26, 2013, 18 UTC (right).

4. ERA-Interim reanalysis of the European Centre for Medium-range Weather Forecasts

Due to the fact that there is a continuous development in the numerical weather prediction models, the use of operational models is not the most straightforward way to go if we would like to intend to study climatological characteristics of any atmospheric phenomena. To solve this problem, several reanalysis projects started applying a frozen version of the model to provide homogeneous quality of the analysis and forecasts.

The European Centre for Medium-Range Weather Forecast (ECMWF) was established in 1975, with the collaboration of 18 European countries (*Woods*, 2006). The ECMWF was among the first centers which create reanalysis database of a longer time period (from 1979 to 1993), it was ERA-15 project in the mid '90s (*Gibson et al.*, 1997). In 2003, ECMWF produced a new longer reanalysis, ERA-40 for the period between 1958 and 2002, with improved spatial resolution and state-of-the-art 3D data-assimilation (*Uppala et al.*, 2005).

The ERA-Interim reanalysis project started in the middle of the first decade of the 21st century. At first, the starting year of this reanalysis was 1989, but it was extended to 1979 (*Dee et al.*, 2011). The ERA-Interim database gets updated each month, with a 2 months delay. Using even more advanced spatial resolution (0.75*0.75 degrees) and the newest 4D-var dataassimilation techniques (as opposed to ERA-40s 3D-var technology) led to a significant increase in the quality of ERA-Interim.

As it was shortly summarized above, the currently available datasets provide state-of-the-art tools for studying cold drops in our region. In addition to model and pressure level fields, several selected predefined isentropic fields are also available from ERA-Interim. It widely supports to extend the potential tools and methods for recognition and investigation of the cold drops.

Beside the importance of the reanalysis datasets, we have to shortly summarize the main characteristics of the current operational models, too. Since June 25, 2013, the deterministic model has contained 137 levels in the vertical and its horizontal resolution is 16×16 kms. Ensemble model contains 62 levels in the vertical and its horizontal resolution is 32×32 kms. In the autumn 2013 vertical resolution of the ensemble model will be increased to 91 levels. The above-mentioned developments have positive impacts on forecasting of the ULLs too. Forecasted meteorological fields are available at standard pressure levels, all model levels, and selected isentropic levels from deterministic and ensemble models.

5. Developments of new methods for statistical and meteorological studies of cold drops

5.1. The horizontal and vertical structure of cold drops

Due to the fact that cold drops relatively rarely occur at any geographical locations, it is not an easy task to get a quite large sample if we plan to study and summarize the typical characteristics of cold drops. For solving the above mentioned problem, we planned to collect cold drops from a larger area, so we needed to develop an objective method for recognition of cold drops. Firstly, several occurrences were collected when cold drops was determined by forecasters in the last ten years. Secondly, the general characteristics of cold drops were determined by applying our newly developed methods. Finally, we plan to extend the sample to investigate cold crops by applying objective recognition algorithm for a 30-year period.

At first, we gathered 70 cases from those submitted by synoptical meteorologists as ULL situations between 2002 and 2011 (*Gaál*, 2012a). These 70 cases were studied daily in 6-hour intervals, so this led us to have 280 different states. At first, a few meteorological parameters as temperature, geopotential, relative humidity, and winds were commonly used at standard pressure levels (850, 700, 500, and 400 hPa), so this approach supported our examination of 3D structure in the atmosphere. Since cold crops have characteristically small horizontal extension, and they rarely occur at each geographical location, we had to choose an area which is large enough for our examinations. The northwestern counterpoint of the area was at latitude N 60° and longitude E 10°, while its southeastern counterpoint was at latitude N 40° and longitude E 40° (*Fig. 3*). *Fig. 3* shows the core positions of the 280 ULLs studied between 2002 and 2011. As we can see, the minimum temperature spots of the cold drops are located mostly west, northwest, and north from Hungary. The purple spots mark the center of the ULLs, and they mostly appear west, north, and northeast from Hungary. All the data used came from the ECMWF MARS database (*Raoult*, 2001).

In the period between 1979 and 2008, we calculated the monthly average temperatures at 500 hPa. We looked at the 30-year average, the minimum and maximum monthly-mean temperatures for each month (*Fig. 4*). Based on the timeline of these 30 years, the amplitude is smaller at 500 hPa (15 °C) than at the surface (22–25 °C). The annual thermal values of the mid-troposphere follow those of the surface with about a month delay. This is in accordance with the fact that the atmosphere gets heated mainly from the direction of the ground, and the effects of the minimum and maximum levels of irradiation show up with an offset in the mid-troposphere. The temperature is clearly lower on days with cold drops than the monthly averages of the 30 years. *Fig. 4* shows the monthly average temperatures (at 500 hPa) from the last 30 years (1979–2008). This is marked by the brown line on the picture, and it stands out quite well that the

temperature of the cold drops (marked in purple) resides below the average monthly data in every case. We can also see that ULLs do not have an absolute temperature threshold.

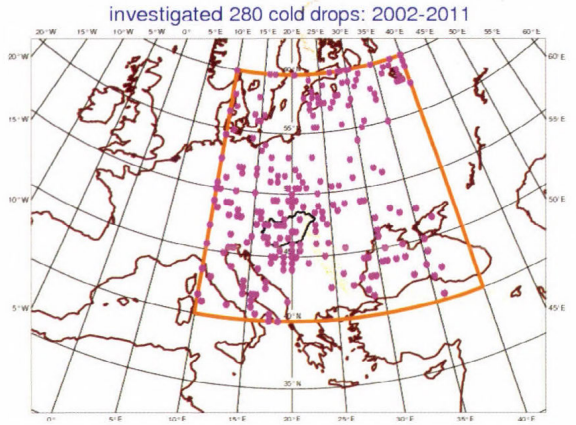

Fig. 3. Centers (local minimums) of the investigated 280 cold drops between 2002 and 2011.

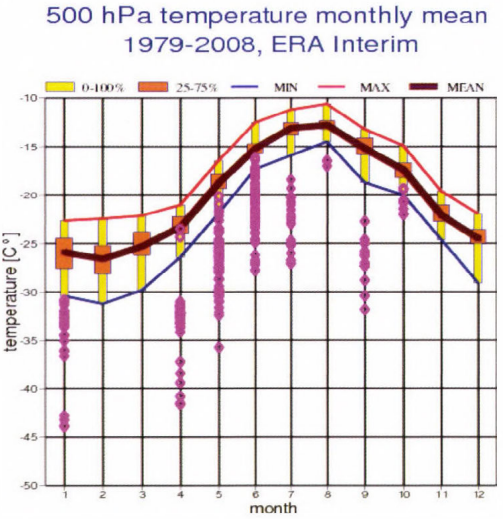

Fig. 4. Monthly means of the 500 hPa temperature (brown line), the coldest monthly means (blue lines), and the warmest monthly means (red lines) in the last 30 years. Purple dots show the investigated minimums of 280 cold drops.

As it has been mentioned, in order to have enough samples for statistical analysis, we have set a goal of generating an algorithm that recognizes ULLs (*Gaál*, 2012a, *Gaál*, 2012b). Firstly, local minima of the temperature fields are

determined at 400, 500, 700, and 850 hPa standard pressure levels. Secondly, we calculated horizontal gradients with several radius around the minima. Lastly, we calculated the axis inclination of cold drops at a few layers in the vertical. In the next few paragraphs, the results are briefly summarized.

Around the local minima, horizontal gradients were calculated within a circle determined by radius, which was changed between 100 and 750 km. Histograms were made from the 280 elements sample for each selected radius and pressure. Typical characteristics of the horizontal structure of ULLs were studied by the investigation of these histograms.

Our results match up with the anticipated structure, which means that gradients are getting lower as we get further away from the core. The unique horizontal temperature structure of cold drops gives us a good way of distinction from the cyclones of temperate zones, which are characterized by a much larger extension and thermal dissymmetry (except for their occluded phase). Another unique trait of cold drops is that their inner core is only clear in the upper troposphere, while in the lower troposphere, it is barely visible (as opposed to those of temperate cyclones). The horizontal thermal gradients are calculated around the core at 100, 250, 500, and 700 km. We also made a histogram of the aforementioned 400, 500, 700, and 850 hPa pressure levels with a gradient frequency of 0.5 °C/100 km (*Fig. 5*).

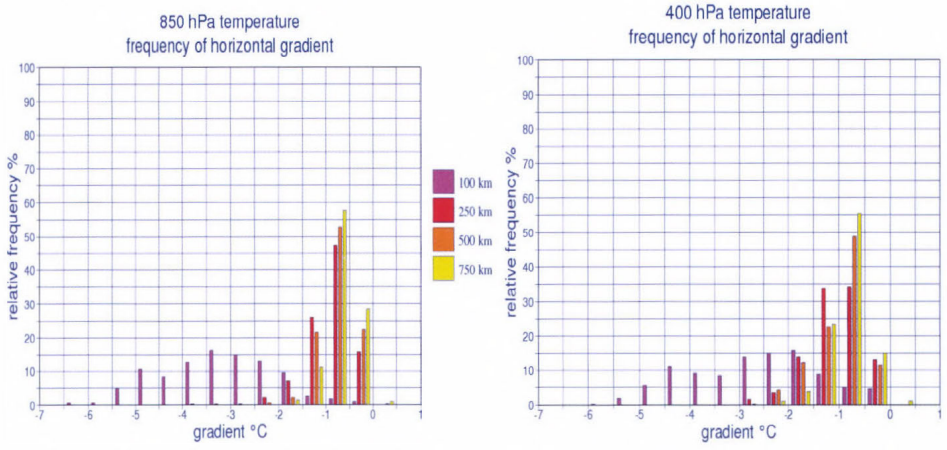

Fig. 5. Histograms of the gradient around the center of cold drops from 100 to 750 km circles at 850 hPa (up) and 400 hPa (down).

Gradients studied at 100 km are usually much larger than those studied at larger areas, and the values of the gradients get lower as the studied area gets larger, since ULLs have a characteristically small horizontal extension. We also made a cross chart showing the gradients calculated at the main isobaric-pairs.

Our calculation of the gradient pairs of 750/100 km, 500/100 km, and 250/100 km allowed us to establish that the gradients calculated on a larger area are always lower than those calculated on a smaller area (*Fig. 6*).

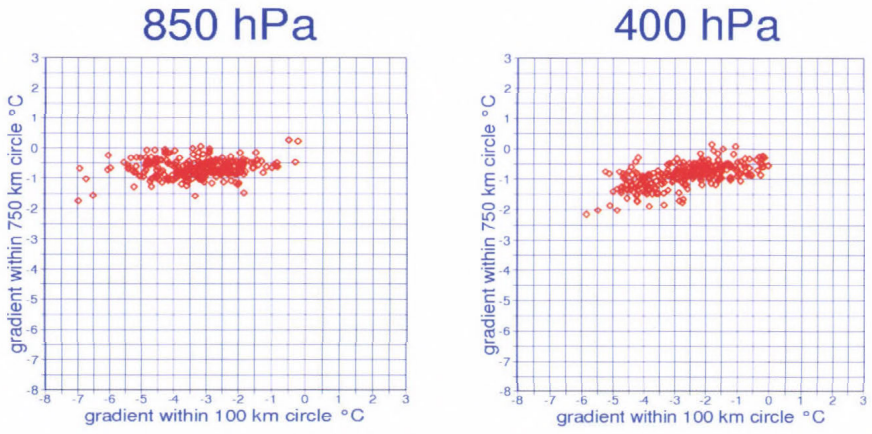

Fig.6. 750/100 km gradient pairs at 850 hPa (left) and 400 hPa (right) pressure levels

We also studied the relative spatial displacement of the cold drops on the mentioned pressure levels, or in other words, the axis inclination of the ULLs. *Fig. 7* shows the axis inclinations at the 400/500 hPa and 400/850 hPa layers. As we can see, the inclination is much lower at 400/500 hPa than at the higher-difference layers, with the axis usually being straight, or leaning towards the east of southeast (which also allows a clear study of the orientations towards different directions).

5.2. Relationship between meteorological fields in cold drops

In this part of the work, we aimed to extensively extend the number of the applied meteorological variables for studying cold drops. Additional variables are 300 hPa wind speed, potential temperature of the 2 potential vorticity unit (2PVU), isentropic potential vorticity at 315 K, horizontal temperature advection at 500 hPa, and wind shear between 850 and 300 hPa. In this paragraph, considering the technical limitations of this paper, only two new meteorological variables are briefly mentioned below.

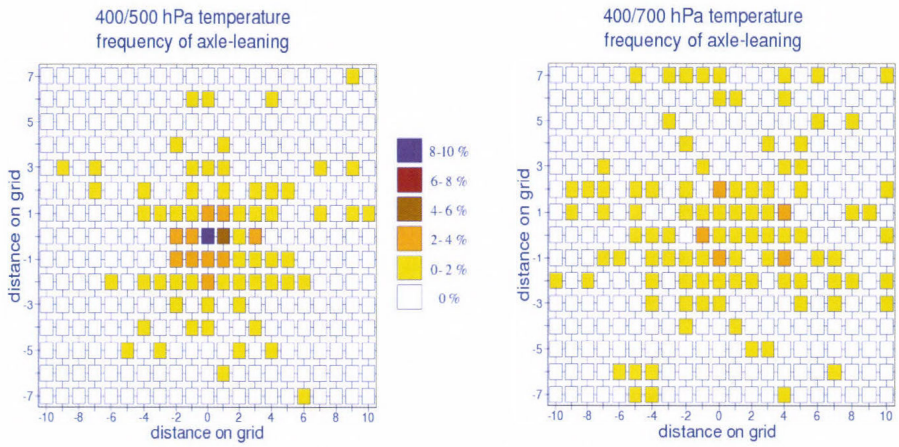

Fig. 7. Frequency of axis leaning at the 400/500 hPa layer (left) and 400/700 hPa layer (right).

Studying connection between cold drops and jet streams we found, that there is a very strong relationship between the position of the cold core and the structure of the jet stream (*Fig. 8*). Before the tear-off phase, the area of the most intensive part of the jet stream quickly moves to south at most part of the wave. At the next step it moves toward to northeast, this change causes a tearing of the drop. In most cases due to further rotation of the cold drop, two intensive parts of the wind shear are visible. One of them appears to the left and the other one appears to the right of the core. In cases of larger cold vortexes, 3 or 4 separated intensive parts of the jet streams can be found quite often. At the final stage of the lifecycle, when the cold drop joins to the main stream, an intensive part of the jet is always found at the eastern part. When the cold drop does not join the main stream it becomes stationary, increased warming at the core is typical, and the intensity of the jet stream decreases.

Besides studying cold drops on standard pressure levels, the investigation could benefit by applying "pv-thinking". The potential vorticity (pv) was introduced by *Ertel* (1942), it was not commonly used until the middle 1980s (*Hoskins et al.*, 1985; *Hoskins*, 1991; *Hoskins*, 1997). The potential vorticity theory is based on the structure of the potential vorticity field. Potential vorticity is the absolute circulation of the air mass between two isentropic surfaces. In the troposphere, its value is usually low. Approaching the tropopause, it sharply rises in the vertical (from 1 PVU to 4 PVU, 1 PVU=10^{-6} $m^2s^{-1}Kkg^{-1}$) with 2PVU being the dynamic tropopause, the subject of our studies.

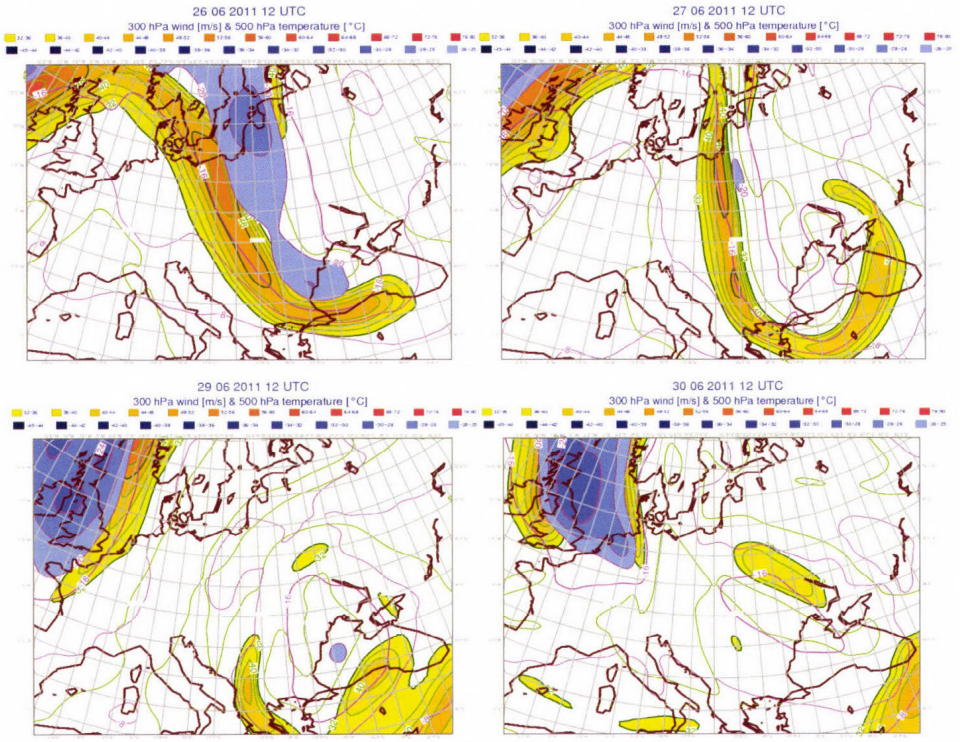

Fig. 8. 300 hPa wind stream and 500 hPa temperature at 12 UTC, June 26, 27, 29, and 30, 2011.

Studying the relationship between the local minimum of the isentropic level of the 2PVU and the position of the cold drops at 500 hPa we found, that these two minimums are typically very close to each other (*Fig. 9*). If there are more than one cold cores, there are consequently more minimums on the 2PVU fields too. We found that in tearing phase there was always a short delay (6–18 h) of the 2PVU field with tailing shape comparing to cold core at 500 hPa.

5.3. New ensemble based graphical products for supporting the recognition of cold drops

Predictability of cold drops of the numerical weather prediction models is not an easy task. It is highly recommended to use ensemble forecasts besides deterministic forecasts. At the Hungarian Meteorological Service, wide range of the ensemble based graphical products, among them EPS plumes, meteograms, spagetthy and probability maps are available in the operational forecasting practice. As a result of our former investigation of the cold drops, it would have

been useful to develop new tools for supporting better recognition and forecast of the cold drops. As in Section 4.1 it was clearly stated, additional meteorological fields besides the 500 hPa temperature can provide very valuable information as far as cold drops are concerned. We developed two new types of ensemble plume diagrams. First one contains three variables: 500 hPa temperature, isentropic potential vorticity at 315 K isentropic level, and potential temperature of the 2PVU surface. The second one contains four variables: in addition to previous three variables, 300 hPa wind speed is also part of the graph.

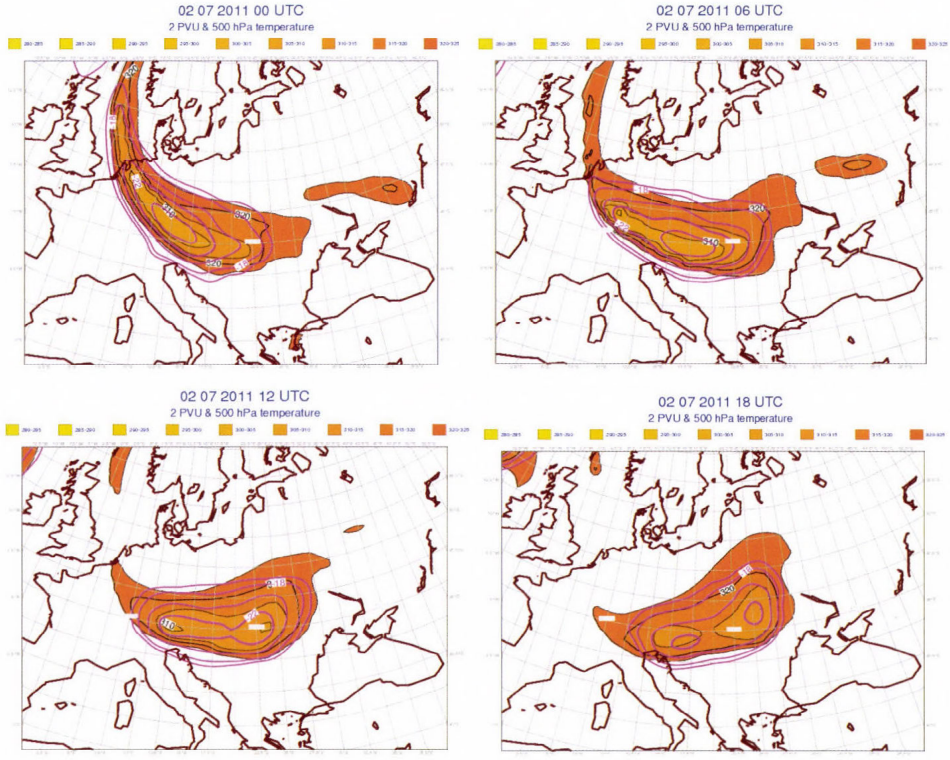

Fig. 9. Potential temperature of the 2PVU surface and the 500 hPa temperature between 00 and 18 UTC, July 2, 2011.

Usefulness of these new tools is demonstrated in a case study. On June 30, 2011, a cold drop pathed across Hungary. *Fig. 10* shows new ensemble plume diagram, ULL occurred on June 30, 2011. We can see the strong U-shape on the second forecast day at 500 hPa, showing a high chance of about 8 degrees drop in temperature in 24 hours and the same amount of rising temperature after the

cold drops passing. *Fig. 10* displays the uncertainty of ULL core forecasting and intensity perfectly.

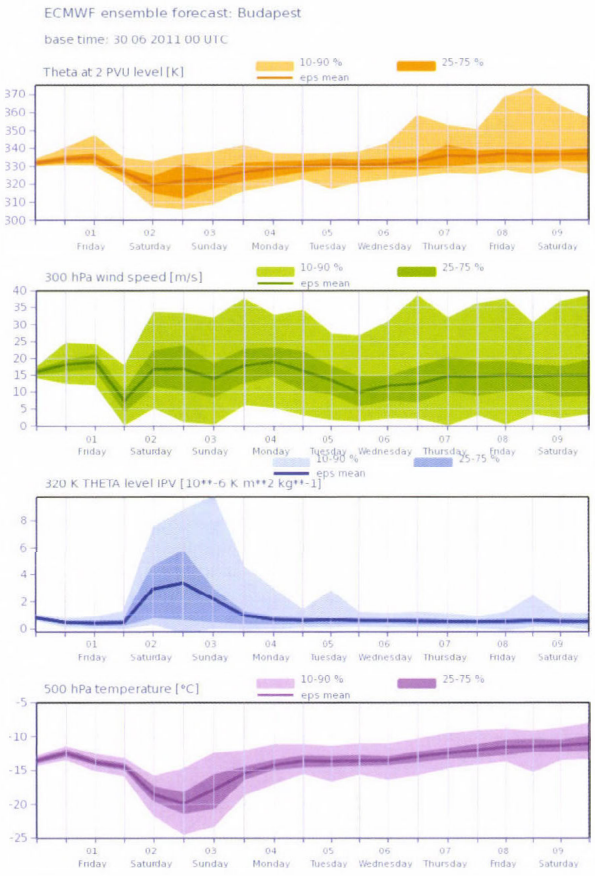

Fig. 10. New ensemble plume diagram, containing potential temperature of the 2PVU surface, isentropic potential vorticity at 315 K isentropic level, wind speed at 300 hPa, and temperature at 500 hPa, for Budapest, model started at 00 UTC, June 30, 2011.

6. Summary, conclusions

Due to the fact that cold drops relatively rarely occur at any geographical locations, it is not an easy task to get quite a large sample if we plan to study and summarize the typical characteristics of cold drops. For solving the above-

mentioned problem, we collected cold drops from a larger area and developed an objective method for recognizing cold drops. Firstly, several situations were collected when cold drops were determined by forecasters in the last ten years. Secondly, the general characteristics of the cold drops were determined by applying our newly developed methods.

Besides determining the general characteristics of the horizontal and vertical structure of cold drops, several new methods were developed for studying and providing guidance for forecasting cold drops. The two most important aspects are the usage of the ensemble forecasts and model forecasts on isentropic levels. At the operational practice, using ensemble forecasts besides the deterministic model can provide very valuable additional information due to the fact that the intensity and position of the cold drops are quite often uncertain. In addition to applying standard pressure level fields, potential temperature of the 2PVU and isentropic vorticity fields are also useful, they could provide more realistic features of the cold drops than we can see from pressure levels. The objective ULL-recognition is required to be able to reliably distinguish cold drops from cyclones. Our results so far make this a very real and reachable goal. One of our plans is to designate the areas with a potential for the formation of cold drops based on ensemble forecasts. In the future, we would like to run further tests with our ULL-recognition algorithm to study the last 30 years of cold drops, and we would also like to experiment more with ULL forecasting.

Acknowledgements: Authors would like to thank *Zoltán Fodor* and *Ákos Horváth* for their motivating suggestions and comments on work. Thanks for anonym reviewers for comments on improving this article.

References

Bodolainé, J.E: 1983: Árhullámok szinoptikai feltételei a Duna és a Tisza vízgyűjtő területén. *OMSZ Hivatalos Kiadványai 56*, Budapest. (In Hungarian)

Bonta, I., Molnár K., and *Vadkerti, F.*, 1989: Zöldár a Tiszán. *Légkör 34(3)*, 8–12. (In Hungarian)

Dee D.P, Uppala S.M., Simmons A.J., Berrisfordm P., Poli P., Kobayshi S., Andrae U., Balmaseda, M.A., Balsamo, G., Bauer P., Bechtold P. Beljaars A.C.M., Van de Ber L., Bidlot J., Bormann N., Delsol C., Dragani R., Fuentes M., Geer A.J., Matricardi M., McNally A.P. Monge-Sanz B.M., Morcrette J.J., Park B.K., Puebey C., Rosnay P., Tavolato C., Thépaut J.N., and *Vitart* F., 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc. 137*, 553–597.

Ertel, H., 1942: Ein neuer hydrodynamischer Wirbelsatz. *Meteorol. Z. 59*, 277–281.

Gaál, N., 2012a*:* Study of cold drops based on ECMWF ERA Interim over Europe *BSc thesis.* Budapest.

Gaál, N., 2012b*:* Study of cold drops based on ECMWF ERA Interim over Europe *Proceedings of the Scientific Student Conference on Environment, 5-7 April 2012*, Veszprém, Hungary.

Gibson, J.K., Kallberg, P., Uppala, S., Hernandez, A., Nomura, A. and *Serrano, E.*, 1997: ERA-15 description, *ERA-15 Project series 1.* ECMWF.

Hoskins, B.J., McIntyre, M. E., and *Robertson, A. W.*, 1985: On the use and significance of isentropic potential vorticity maps *Q. J. R. Meteorol. Soc. 111*, 877–946.

Hoskins, B.J., 1991: Towards a PV - Theta view of the general circulation; *Tellus 43 AB*, 27–35.

Hoskins, B.J., 1997: A potential vorticity view of synoptic development. *Meteorol. Appl. 4*, 325–347.

Isaksen, L., Haseler, J., Buizza, R., and *Leutbecher, M.,* 2010: The new ensemble of Data Assimilations, *ECMWF Newsletter 123*, 17–21.

Horváth, Á., 2007: Components of the convection in the atmosphere (in Hungarian). In (ed. *Horváth, Á.*) Atmosheric convection, *Hungarian Meteorological Service*. Budapest. 4–17.

Kurz, M., 1990: Synoptische Meteorologie. *Verlag Deutscher Wetterdienst*, Offenbach.

León F.M., 2003: Las gotas frías / DANAS: Ideas y conceptos básicos. Servicio de Técnicas de Análisis y Predicción, INM. Instituto Nacional de Meteorología. (In Spain)

Nieto, R., Gimeno, L., Torpe, L, Ribera P., Gallego, P., Garcia-Herrera, R., Garcia, J.A, Nunez, M., Redano, A., and *Lorente, J.,* 2005: Climatological features of cutoff low systems in the Northern Hemisphere. *J. Climate 18,* 3085–3012.

Pelly, J.L. and *Hoskins, B.J.,* 2003: A New Perspective on Blocking. *J. Atmos. Sci. 60*, 743–755.

Peltonen, T., 1963: A Case Study of an Intense Upper Cyclone over Eastern and Northern Europe in November 1959. *Geophisica 8*, Helsinki, 225–251.

Raoult, B., 2001: MARS on the web: a virtual tour. *ECMWF Newsletter 90*, 9–17.

Rex, D.F., 1950a: Blocking action in the middle troposphere and its effect upon regional climate. Part I. An aerological study of blocking action. *Tellus 2*, 196–211.

Rex, D.F., 1950b: Blocking action in the middle troposphere and its effect upon regional climate. Part II. The climatology of blocking action. *Tellus 2*, 275–301.

Scherhag, .R, 1948: Neue Methoden der Wetteranalyse und Wetterprognose. Springer, Berlin.

Simmons, A., Uppala, S., Dee, D., and *Kobayashi, S.,* 2007: ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter 110*, 25–35.

Szépszó, G., 2003: A 80-as és 90-es évek árhullámainak szinoptikus-klimatológiai értékelése. Master thesis. Eötvös Loránd University, Budapest. (In Hungarian)

Uppala, S.M., Kallberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.F., Morcrette, J.J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., and *Woollen, J.,* 2005: The ERA-40 Reanalysis. *Quart. J. Roy. Met. Soc. 131*, 2961–3013.

ZAMG, 2007: Manual of synoptical and satellite meteorology: Conceptual models and cases studies. http://www.zamg.ac.at/docu/Manual/

Woods, A., 2006: Medium-Range Weather Prediction The European Approach. Springer.

NEWS

IN MEMORIAM JEAN-FRANÇOIS GELEYN (1950-2015)

Jean-François Geleyn passed away on January 8, 2015. We knew that Jean-François was struggling with his health in the last years, but still the incoming news was devastating for everybody in the Numerical Weather Prediction (NWP) community. We lost a leader in NWP in general and our leader in the ALADIN cooperation in particular. Jean-François had enormous influence on the progress of NWP in Central and Eastern Europe certainly including also Hungary. He had special ties with Hungary as honorary member of the Hungarian Meteorological Society (since 2000) and a member of the Editorial Board of "Időjárás". He paid 12 visits to Hungary between 1992 and 2011 (his last talk in Hungary can be seen at http://www.mettars.hu/wp-content/uploads/2015/01/Eloadas110328.pdf).

Jean-François Geleyn was the main initiator of the ALADIN international cooperation in 1990 with the proposal of the support programme of Météo France towards several Central and Eastern European countries. Hungary joined the cooperation among the first countries and already took part (by Dezső Dévényi) in the feasibility study of the project in March, 1991 in Paris. The development work of the ALADIN limited-area model started in Toulouse in September, 1991. Jean-François was leading the group in charge of the development of the ALADIN model. He was not only our scientific supervisor. He was also very supportive to help the newly arriving Météo France "stagiaires" (the French word for students) from Prague, Budapest, Vienna or Bucharest to find their way into their new lives in Toulouse. We remember very well, he was waiting for us at "Gare Matabiau" at our first arrival to Toulouse. He helped us to settle down in our new temporary home in France, he was showing us around in Toulouse and at Météo France, he was jogging with us, when we prepared for our marathon along the Canal du Midi and he was keen to join our "stagiaire" parties at the Météo France campus. His attitude and helpfulness made possible to "cement together" the "students" in Toulouse inside and outside work. This strong cohesion among the ALADIN developers was one of the secrets of the success of the ALADIN project. This made possible that the researchers working together on ALADIN was not simply forming a scientific team, but also a friendly and supportive company, which is spanning much farther than work on Numerical Weather Prediction. Based on all these scientific and personal achievements it was not accidental that the HIRLAM cooperation decided (around 10 years ago) to join forces with the ALADIN partners and further develop the ALADIN system together (they refer to it as HARMONIE). Now the ALADIN and HIRLAM communities together are forming the strongest NWP limited-area consortium in Europe.

Jean-François's main area of interest was physical parameterisations of the NWP models, but he was extremely knowledgeable at every aspect of numerical modelling. He had an excellent overview of the entire NWP and he was a very fast thinker making it almost impossible to follow him real time. He had a firm long-term vision on NWP, he was stubbornly insisting on his views and he was tirelessly working on its realisation. Basically, he had never had a rest, he devoted his entire life to NWP and to the ALADIN cooperation.

We all know that the landscape of numerical weather forecasting modelling in Europe would be very different without the contributions of Jean-François. We owe him enormous professional and personal gratitude for all what he had done for meteorology. We will always remember him as a scientist and friend who was always supporting us and our meteorological community. We will miss Jean-François with sorrow, but we will preserve his memory and put forward his legacy in Numerical Weather Prediction.

András Horányi and *Gábor Radnóti*

INSTRUCTIONS TO AUTHORS OF *IDŐJÁRÁS*

The purpose of the journal is to publish papers in any field of meteorology and atmosphere related scientific areas. These may be

- research papers on new results of scientific investigations,
- critical review articles summarizing the current state of art of a certain topic,
- short contributions dealing with a particular question.

Some issues contain "News" and "Book review", therefore, such contributions are also welcome. The papers must be in American English and should be checked by a native speaker if necessary.

Authors are requested to send their manuscripts to

Editor-in Chief of IDŐJÁRÁS
P.O. Box 38, H-1525 Budapest, Hungary
E-mail: journal.idojaras@met.hu

including all illustrations. MS Word format is preferred in electronic submission. Papers will then be reviewed normally by two independent referees, who remain unidentified for the author(s). The Editor-in-Chief will inform the author(s) whether or not the paper is acceptable for publication, and what modifications, if any, are necessary.

Please, follow the order given below when typing manuscripts.

Title page: should consist of the title, the name(s) of the author(s), their affiliation(s) including full postal and e-mail address(es). In case of more than one author, the corresponding author must be identified.

Abstract: should contain the purpose, the applied data and methods as well as the basic conclusion(s) of the paper.

Key-words: must be included (from 5 to 10) to help to classify the topic.

Text: has to be typed in single spacing on an A4 size paper using 14 pt Times New Roman font if possible. Use of S.I. units are expected, and the use of negative exponent is preferred to fractional sign. Mathematical formulae are expected to be as simple as possible and numbered in parentheses at the right margin.

All publications cited in the text should be presented in the *list of references*, arranged in alphabetical order. For an article: name(s) of author(s) in Italics, year, title of article, name of journal, volume, number (the latter two in Italics) and pages. E.g., *Nathan, K.K.,* 1986: A note on the relationship between photosynthetically active radiation and cloud amount. *Időjárás 90,* 10-13. For a book: name(s) of author(s), year, title of the book (all in Italics except the year), publisher and place of publication. E.g., *Junge, C.E.,* 1963: *Air Chemistry and Radioactivity.* Academic Press, New York and London. Reference in the text should contain the name(s) of the author(s) in Italics and year of publication. E.g., in the case of one author: *Miller* (1989); in the case of two authors: *Gamov* and *Cleveland* (1973); and if there are more than two authors: *Smith et al.* (1990). If the name of the author cannot be fitted into the text: *(Miller,* 1989); etc. When referring papers published in the same year by the same author, letters a, b, c, etc. should follow the year of publication.

Tables should be marked by Arabic numbers and printed in separate sheets with their numbers and legends given below them. Avoid too lengthy or complicated tables, or tables duplicating results given in other form in the manuscript (e.g., graphs).

Figures should also be marked with Arabic numbers and printed in black and white or color (under special arrangement) in separate sheets with their numbers and captions given below them. JPG, TIF, GIF, BMP or PNG formats should be used for electronic artwork submission.

More information for authors is available:
journal.idojaras@met.hu

Published by the Hungarian Meteorological Service

Budapest, Hungary

INDEX 26 361

HU ISSN 0324-6329