

Wissenschaftliche Mitteilungen

aus dem
Institut für Meteorologie der Universität Leipzig



ISBN 3-9807315-1-0

Meteorologische Arbeiten aus Leipzig (VI)

Hrsg.: A. Raabe und K. Arnold

Leipzig 2001

Band 22

Inhaltsverzeichnis

Wissenschaftliche Mitteilungen aus dem Institut für Meteorologie der Universität Leipzig und dem Institut für Troposphärenforschung e.V. Leipzig, Band 22

	Seite
Concepts for coupling hydrological and meteorological models N. Mölders	1
Sensitivity studies with a surface and channel runoff module coupled to a mesoscale atmospheric model N. Mölders und W. Rühaak	16
Numerical solution of the Stochastic Collection Equation – Comparison of the Linear Discrete Method and the Method of Moments M. Simmel	26
Berücksichtigung atmosphärischer Schallausbreitungsbedingungen beim Lärmschutz A. Ziemann, K. Arnold und A. Raabe	35
Acoustic tomography as a method to characterize measuring sites A. Ziemann, K. Arnold und A. Raabe	50
Acoustic tomography in comparison to in-situ temperature and wind measurements K. Arnold, A. Ziemann und A. Raabe	60
Messung der Lufttemperatur und Luftfeuchte unter Berücksichtigung des Strahlungsfehlers auf einem Versuchsgelände S. Schienbein und K. Arnold	69
Zur Schätzung von Häufigkeitstrends von extremen Wetter- und Klimaereignissen M. Mudelsee, M. Börngern und G. Tetzlaff	78
Datentransfer und deren Wirkung in der spektralen Varianzanalyse geophysikalischer Meßreihen H.-J. Schönfeldt und S. von Löwis	89
A Possible Connection of Midlatitude Mesosphere/Lower Thermosphere Zonal Winds and the Southern Oscillation Ch. Jacobi und D. Kürschner	98
Einfluß des winterlichen stratosphärischen Polarwirbels auf die zonale Symmetrie des Windfeldes in der oberen Mesosphäre und unteren Thermosphäre simuliert mit dem COMMA-Modell M. Lange und Ch. Jacobi	111
Das lokale Optimierungsproblem H. Heinrich und W. Metz	121
Singular vectors of Eady-models with $\beta \neq 0$ and $q' = 0$ R. Faulwetter und W. Metz	137

Concepts for coupling hydrological and meteorological models

Nicole Mölders

Summary: Earth system modeling, climate modeling, water resource research as well as integrated modeling (e.g., climate impact studies) require the coupling of hydrological and meteorological models. The paper presents recent concepts on such a coupling. It points out the difficulties to be solved, and provides a brief overview on recently realized couplings. Furthermore, a concept of a hydrometeorological module to couple hydrological and meteorological models is introduced.

Zusammenfassung: Wasserressourcenforschung, Erdsystem- und Klimamodellierung sowie integrierte Modellierung (z.B. Klimafolgenforschung) erfordern das Koppeln von hydrologischen und meteorologischen Modellen. Dieser Artikel präsentiert Konzepte für eine solche Kopplung. Er zeigt die zu lösenden Schwierigkeiten auf und gibt einen kurzen Überblick über bisher realisierte Kopplungen. Ferner stellt er ein Konzept für einen hydrometeorologischen Moduls zur Kopplung von hydrologischen mit meteorologischen Modellen vor.

1. Introduction

In nowadays there are several reasons and interests to couple hydrological and meteorological models. Climate impact studies, planing of water resources or landscapes as well as ecosystem modeling require to describe the water cycle in a closed manner. Due to the different aspects of the water cycle that are of major or minor interest to hydrologists, meteorologists or climatologists only those processes are considered in detail that are relevant for the specific application of interest while other aspects are neglected or crudely simplified. As a consequence the water cycle is not simulated in a closed manner. The simplifications, however, can provide difficulties in the description of the water cycle even in fields of non-interdisciplinary modeling. In traditional numerical weather prediction (NWP) models, for example, the neglecting of lateral soil water movements and surface runoff may yield to an underestimation of soil moisture in river valleys and an overestimation of soil moisture in the nearby mountainous regions, which usually receive more precipitation, if the simulated soil moisture distribution of the previous day is used to initialize the soil moisture distribution for the next forecast (e.g., Müller et al. 1995). Consequently, the local recycling of previous precipitation may be predicted incorrectly which may lead to wrong forecasts of convective clouds, showers, and thunderstorms. In weather forecasting or climate modeling, for instance, a pre-given constant water table may provide errors in the water supply to the atmosphere. In nature, namely, the water table will rise during long-lasting precipitation episodes in areas of flat water tables, while it will sink or even decouple from soil moisture during long-lasting drought episodes. Evidently, during droughts, the lowered ground water table may contribute to the duration of the drought because evapotranspiration, ET, is reduced (Fig. 1). On the contrary, long-lasting extreme precipitation events may trigger their persistence by recycling of previous precipitation because the high water table yields in enough water available for evapotranspiration (Fig. 1). In hydrological modeling, for instance, neglecting the feedback of increased soil moisture on evapotranspiration, cloud- and precipitation-formation can lead to mispredictions of flood intensity, because runoff of an individual river catchment depends on precipitation and evapotranspiration within the basin (e.g., Liston et al. 1994). Water storage within the river basins among others depends on soil type, soil depth, surface heterogeneity

and vegetation cycle (Miller et al. 1994) where the latter again depends on microclimatological conditions which are not simulated. In general circulation models (GCM), for example, river runoff provides the critical link for returning water from continents to the ocean (Miller et al. 1994). Although at any given time rivers hold only a fraction of the world's total water river runoff is an important input value for ocean models because freshwater flow affects the thermohaline circulation of the ocean.

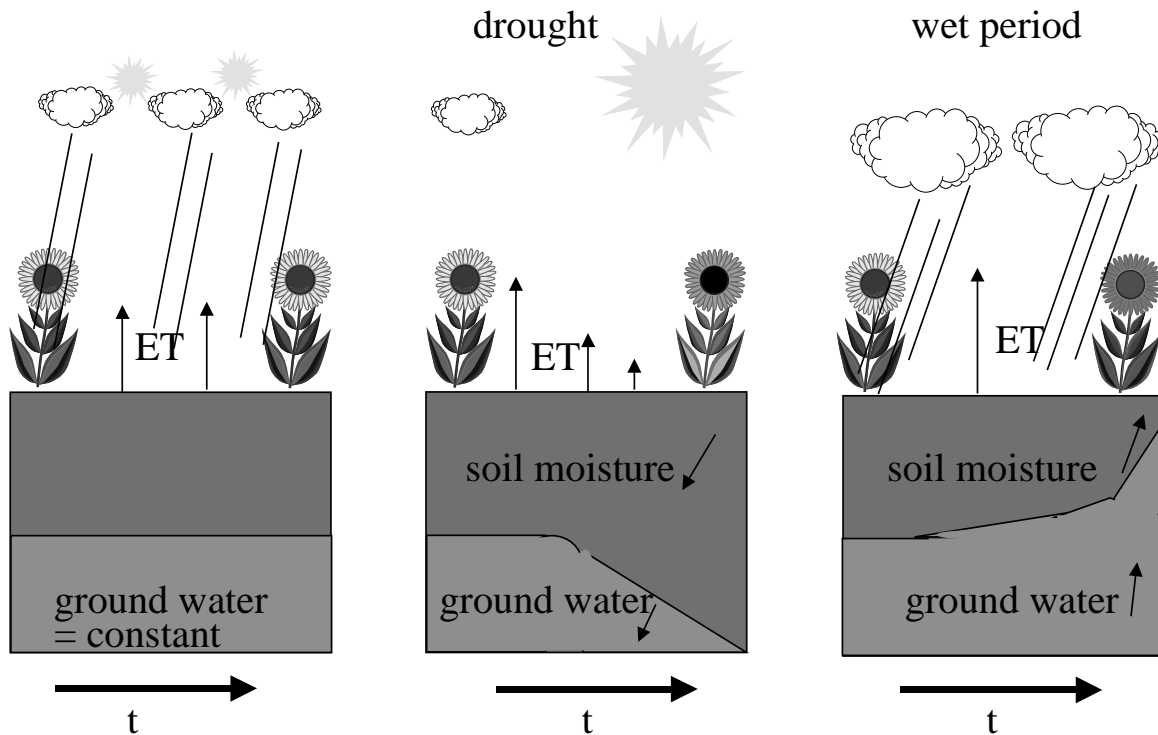


Fig. 1. Schematic illustration of the impact of a flat water table on atmospheric water fluxes.

To more appropriately consider the interaction of the atmospheric and land-based path of the water cycle a closed description of the water cycle is an urgent need. In this paper, recently published concepts to model the water cycle are evaluated and a concept to couple hydrological and atmospheric models by use of a hydrometeorological module are presented.

2. The coupling problem

As pointed out above, meteorologists and hydrologists have a common interest in land-surface modeling, but both by tradition and purpose, the land surface parameterizations applied in hydrological and atmospheric models differ strongly (e.g., Graham and Bergström 2000). Hydrologic models require a precise partitioning of precipitation into evapotranspiration, infiltration, interception, retention, and runoff to determine the water fluxes within the soil and ground water recharge, i.e., the water balance is of main interest. Meteorological models need a precise partitioning of precipitation between the aforementioned processes to determine the partitioning of incoming radiation between soil heat flux and the turbulent fluxes of sensible and latent heat, i.e., they additionally need the energy balance. The water balance is often very simplified by using a bucket model or a force restore method.

		Microscale				Mesoscale			Macroscale			
		Microscale		Mesoscale			Macroscale				L T	
		δ	γ	β	α	γ	β	α	β	α		
30 a											Climatological scale	
1 a												
1 m												
1 d							thunderstorms	Squall line	cloud cluster	fronts	synoptic scale	
1 h							fog					
1							convection				meso scale	
min							Rain gauge measurements					
1 s							Turbulent water fluxes				microscale	
							cloud microphysics					
		2 mm	2 m	20 m	200 m	2 km	20 km	200 km	2000 km	20000 km		

Fig. 2. Schematic view of spatial (L) and temporal (T) scales in hydrological and meteorological models. The upper and lower termin of scales address the hydrological and meteorological models, respectively. Note that hydrological models used in flood forecasting or warning systems have typical application times of several days to a week with time resolutions of one to several hours, while water balance models address the climatological scale with time resolutions of one day. Macroscale hydrological models range from synoptic to climatological scale with time resolutions of 10 to 30 minutes. See text for further discussion.

Other differences between hydrological and meteorological models are related to the typical scales of processes to be simulated (Fig. 2). Therefore, the temporal and spatial resolutions of these models differ strongly. Meteorological models usually apply square grid cells as their horizontal unit of area. Despite of increasing computational power decreasing of these grid cells is limited by the range of validity of the assumptions made in the parameterizations (e.g., the fetch-conditions from which the parameterizations of the surface fluxes are derived lead to a requirement of a relation of 1:100 for the ratio of the vertical to horizontal grid resolution), and the limited availability of initial data. Hydrological models have to represent adequately the river basin and its sub-basins which consequently define the model dimensions. Thus, especially in complex terrain, the grid resolution may be very fine. Decreasing the horizontal resolution of hydrological models, for instance to match the resolution of meteorological models, provides the difficulty how to represent the (then subgrid-scale) heterogeneity of land surface properties, such as topography, geology, and vegetation cover. As pointed out before these characteristics, among other things, determine the runoff and storage of the basin.

The solution of the energy balance, which is of essential need in meteorological models, demands time steps of several minutes, while a lot of hydrological models are often satisfied with daily time steps (e.g., Graham and Bergström 2000). Meteorological models usually neglect lateral flows of soil water, surface runoff, the transport of water by river flow as well as the re- and discharge of the groundwater storage, because these processes are slow as compared to typical simulation time scales of meteorological models. As aforementioned,

these processes, however, are important in hydrological models and may be of great relevance in climate modeling as well as in weather forecasting (Mölders and Raabe 1997). Thus, one main task in coupling hydrological and meteorological models is to overcome these temporal and spatial gaps.

3. Concepts

3.1 Parameterizations

Being aware of the problematic of an unclosed water cycle and the difficulties in coupling hydrological and meteorological models caused by the different spatial and temporal scales, recently, several authors (e.g., Kuhl and Miller 1992, Marengo et al. 1994, Miller et al. 1994, Sausen et al. 1994, Hagemann and Dümenil 1998) suggested parameterizations of different complexity to directly parameterize runoff in GCMs. Kuhl and Miller (1992) suggested a simple scheme wherein all runoff within a river drainage basin instantaneously reaches the river mouth. Although the global runoff agreed well with the observed runoff, great errors occurred locally. Miller et al. (1994) introduced a river model that allows the excess water at the surface calculated by a GCM to run off into the river within a continental grid cell. The direction and speed of flow is either constant or depends on topography gradient. Sausen et al. (1994) proposed a one-parameter model that represented each grid cell by a two-dimensional linear reservoir with different retention coefficients for the flows in the east, west, north, and south direction. These coefficients depend on the orography and grid size. This approach does not distinguish the different types of flow processes. Thus, in an attempt to improve this approach Hagemann and Dümenil (1998) introduced a global parameterization wherein the cascade of overland and river flow is realized by equal linear reservoirs, and baseflow is considered by a one parameter model. The corresponding retention coefficients depend on topography gradient between two neighbored grid cells and on the grid cell size.

3.2 Direct coupling by data exchange

Despite of the temporal and spatial gaps between the resolutions of hydrological and meteorological models, recently, several different methods have been developed and tested to couple hydrological and meteorological models directly.

3.2.1 One-way-coupling

The simplest version to “couple” hydrological and meteorological models is the one-way coupling wherein the hydrological model is driven by the meteorological model. There exists no feedback of the hydrological model results to the simulation of the meteorological model. It is obvious that, in the case, of large floods, applying a one-way coupling may lead to false water supply to the atmosphere which may result in a wrong cloud and precipitation distribution (e.g., Mölders 1999). From the practical point of view a one-way coupling is not advantageous for flood forecasting, because the errors in predicted precipitation pattern and intensity propagate in the hydrological model (e.g., Mölders et al. 1999a).

3.2.2 Two-way-coupling

In a two-way coupling, a feedback between the atmospheric and hydrological model results exists. This feedback depends on the degree of coupling. Mölders and Raabe (1997) as well as Mölders et al. (1999a) coupled a conceptual hydrological model (NASMO, precipitation runoff model) and a meteorological model (GESIMA, Geesthacht’s simulation model of the atmosphere) in a two-way mode. In doing so, the hydrologic processes of the river catchment (translation, retention, lateral discharge) are considered in the meteorological model which itself drives the hydrological model by predicted evapotranspiration and precipitation. The

quantities provided by the hydrological model served to modify soil moisture in the meteorological model. The different scales of the models were bridged by aggregation of lateral flow and runoff determined by the hydrological model and disaggregation of evapotranspiration, precipitation as well as soil wetness in the meteorological model, i.e., the module designed to couple the models allows to heterogenize precipitation and to consider subgrid-scale surface processes in the meteorological model. A main problem in this way of coupling was that soil wetness, vegetation- and soil-type as well as topography differ between the models because of the conceptual approaches applied in the hydrological model.

To avoid such discrepancies in the surface characteristics of the coupled modeling system, so-called distributed or the lumped integral models should be used as hydrological models (Todini 1988) and GCMs, NWP or regional-climate models should be applied for the atmospheric part. Note that NWP models belong to the group of mesoscale-meteorological models. In the coupling, the meteorological model should be nested into a global model to allow long-term simulations. In order to derive suitable parameterizations (quasi) distributed differential hydrological models should be used which are based on the fundamental laws of hydro-dynamics (e.g., MIKESHE; Abbot et al. 1986, Refsgrad et al. 1992) and that also include hydro-thermodynamic effects like soil frost.

3.3 A hydrometeorological module as intersection

According to the first experiences with loosely coupled hydrological and meteorological models it seems to be necessary that coupled hydrological and meteorological models should become a hydrometeorological model, i.e., one model, to avoid that processes are differently parameterized in the hydrological and meteorological model parts (e.g., Mölders and Raabe 1997, Mölders et al. 1999a). Such a realization would save computer time and man-power, but needs interdisciplinary cooperation. It guarantees that the landscapes, soil processes, and the processes at the earth-atmosphere interface are identical in both model parts and provides a closed and such more appropriate description of the water-cycle than the single models or loosely coupled models.

Recently, some developments into the direction of a hydrometeorological module as an intersection to couple hydrological and meteorological models were carried out. TOPMODEL (Beven and Kirkby 1979) was coupled with SVAT (soil vegetation atmosphere transfer schemes), for example by Famiglietti and Wood (1991), Band et al. (1993), Stieglitz et al (1997). Walko et al. (2000) coupled TOPMODEL into a regional atmospheric modeling system (RAMS). Recent state-of-the-art land surface models (LSM) like OSULSM (Oregon State University land surface model, Chen et al. 1996), HTSVS (hydro-thermodynamic soil vegetation scheme, Kramm et al. 1996, Mölders et al. 1999b, Mölders and Rühaak 2000), SEWAB (Mengelkamp et al. 2000) that are used in mesoscale meteorological models address both hydrological and meteorological aspects. Some LSMs also include a component to consider surface runoff and channel flow by using the unit hydrograph (e.g., SEWAB) or the St.-Vernant-equation (e.g., HTSVS).

In general, in a coupled hydrological and meteorological model system, the hydrological model should calculate the ground water, horizontal soil water fluxes, and runoff. The meteorological model should determine the fields of wind, pressure, air temperature, humidity, radiation, cloudiness, and precipitation. In realizing the coupling, the soil-vegetation module of the hydrological or meteorological model should be the common part of data exchange and use. It serves as the upper boundary condition for the hydrological model and the lower boundary condition of the meteorological models. The hydrometeorological module has to fulfill the requests of both the hydrological and meteorological models in the highest degree of accuracy and physical process details required by one of the models. In the following, the special needs of atmospheric and hydrological models to simulate the boundary

conditions are discussed to elucidate which physical processes are to be considered in the hydrometeorological module with which accuracy.

The hydrometeorological module has to be called with the time step of the meteorological model. The hydrological model can be called at coarser time steps than the meteorological model. It has to be examined whether the usual time step of the hydrological model is sufficient or whether a shorter time step delivers better results for the coupled system.

3.3.1 Representation of finescale variations in surface characteristics and conditions

As pointed out above, the heterogeneity of surface characteristics is of great importance in hydrological modeling. It also plays an important role in meteorological modeling. Friedrich et al. (2000) examined the influence of surface heterogeneity on spatial distribution, temporal development, and on the domain-average of the ratio between sensible and latent heat-flux (Bowen-ratio) for synthetic landscapes of differing degrees of surface heterogeneity applying a mesoscale meteorological model. Their results substantiate that land-surface distributions will non-linearly influence the Bowen-ratio if patches of equal type exceed a certain size and that the surface type dominating a landscape does not necessarily determine the mean Bowen-ratio representative for this area. Thus, when applying the dominant surface type as the representative one for an area of a grid-cell, the margin of error in the Bowen-ratio depends on the horizontal resolution of the model (or on available data).

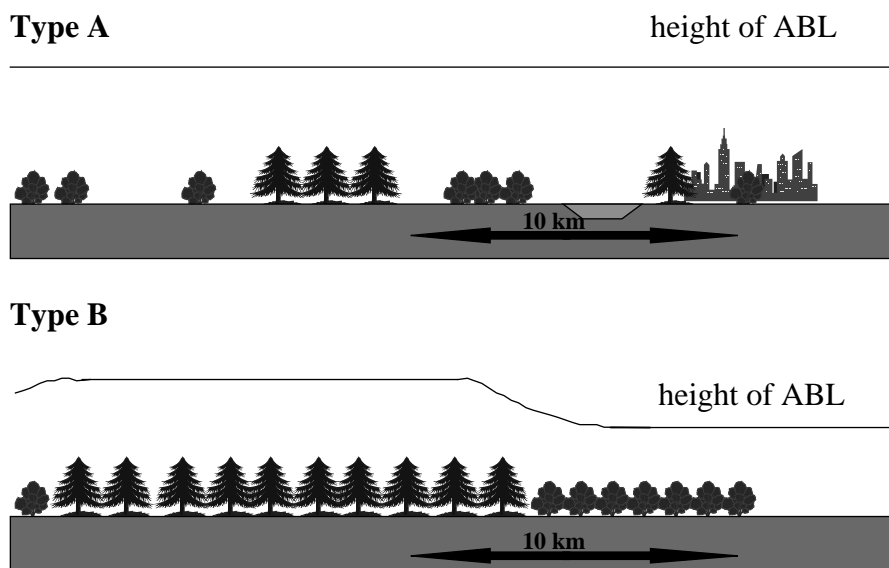


Fig. 3. Schematic view of the affect of mesoscale surface heterogeneity on the height of the ABL (modified after Shuttleworth 1988). Type A and B refer to the classification of heterogeneity type introduced by Shuttleworth (1988).

Therefore, the heterogeneity on the microscale, which is of relevance for the near-surface stratification of the atmosphere (stability) and the atmospheric fluxes of sensible and latent heat, as well as the heterogeneity on the mesoscale, which affect the height of the atmospheric boundary layer (ABL; Fig. 3), vertical mixing, and possible cloud location, should be taken into account. For the reasons discussed above, a special feature of the hydrometeorological module should be its ability to represent finescale variations in surface characteristics, such as

terrain slope, vegetation type, soil type and moisture or water bodies, which often vary considerably over short distances (e.g., Fig. 5).

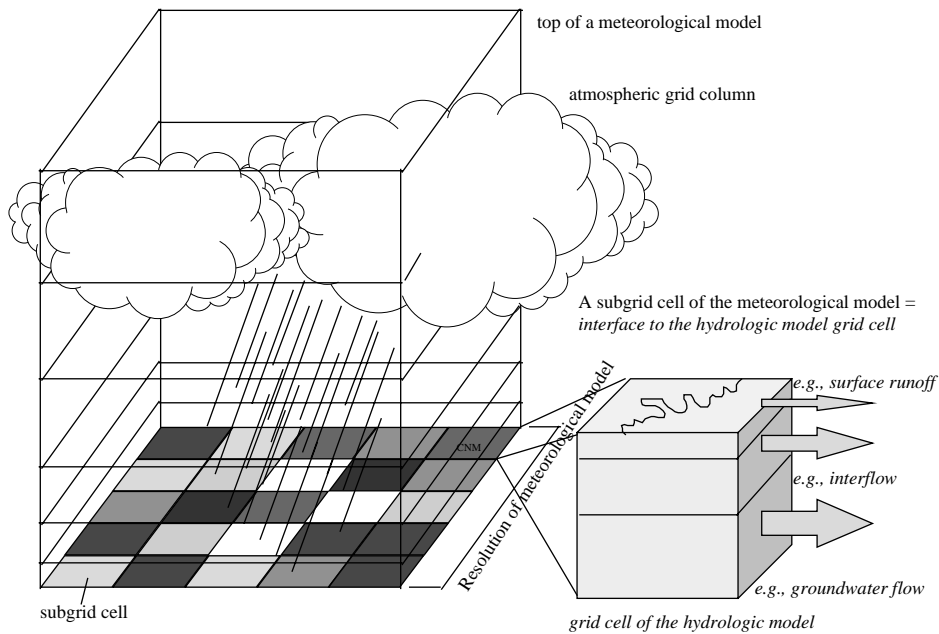


Fig. 4 Schematic view of the downscaling applied in the meteorological model. The different grey levels represent different types of landuse (after Mölders et al. 1999a).

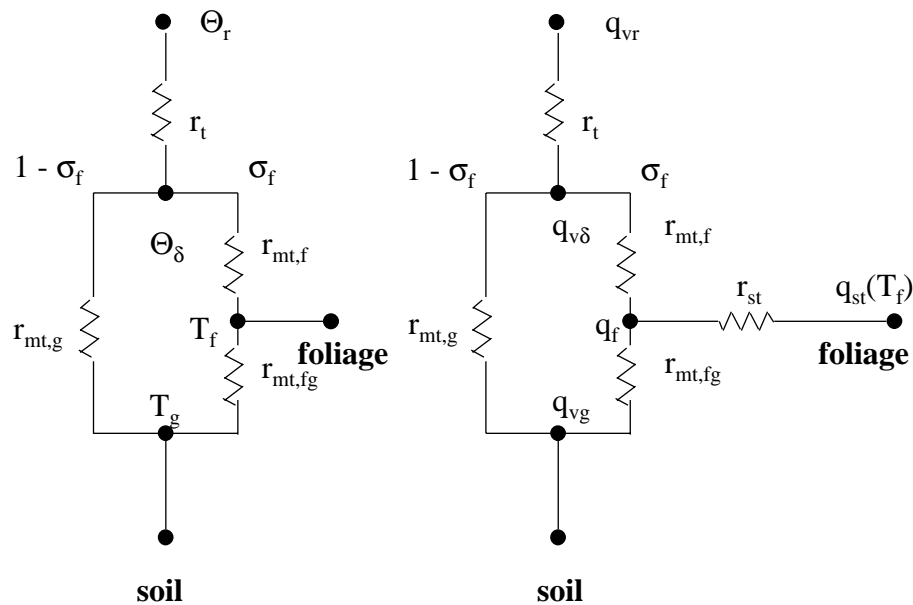


Fig. 5. Schematic view of a resistance network for the mixture approach assuming a dry vegetation shielding the ground by σ percent given in values between 0 and 1 (modified after Kramm et al. 1996). Here, r_{mt} , r_t , r_{st} , T , q_{st} , q_v , and Θ stand for the molecular turbulent resistance, the turbulent resistance, the stomatal resistance, temperature, specific humidity in the stomata, specific humidity, and the potential temperature where the subscripts f, δ , and g stand for the foliage, the reference height and ground, respectively.

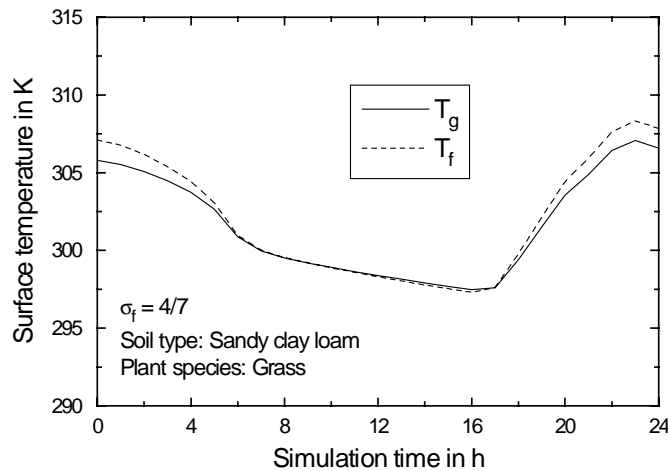


Fig. 6. Temporal development of foliage and ground temperature.

The heterogeneity on the microscale should be included by a mixture approach (e.g., Fig. 5) to consider simultaneously at least bare soil and/or vegetation within one grid-cell (e.g., Deardorff 1978). The surface temperatures of the foliage and the soil as well as their diurnal course, namely, may differ strongly (Fig. 6).

The heterogeneity on the mesoscale can be considered by some type of mosaic approach (e.g., Avissar and Pielke 1989, Seth et al. 1994, Mölders et al. 1996).

Recently, several different strategies have been developed to parameterize subgrid-scale surface heterogeneity, for instance, by averaging surface properties (e.g., Lhomme 1992, Dolman 1992), or by statistical-dynamic approaches (e.g., Wetzal and Chang 1988, Entekhabi and Eagleson 1989). Computationally more expensive procedures to consider patchy surface properties are the mosaic approach (Avissar and Pielke 1989), the explicit subgrid strategy (Seth et al. 1994), or the mixture strategy, wherein for the different surface types tightly coupled energy balances are determined (e.g., Sellers et al. 1986, Dickinson et al. 1986, Kramm et al. 1994). Several authors comparing the results provided by simulations with and without consideration of subgrid-scale surface heterogeneity found that for very patchy surfaces large differences in the predicted fluxes can occur (e.g., Avissar and Pielke 1989, Seth et al. 1994, Mölders and Raabe 1997). A review on methods to treat heterogeneity is given by Giorgi and Avissar (1997).

The difficulty to bridge the spatial scales by aggregation/disaggregation can be addressed within the framework of considering subgrid-scale heterogeneity, i.e., the heterogeneity on the mesoscale. First, parameterizations to downscale hydrologically relevant quantities provided by the meteorological model are required to utilize evapotranspiration and precipitation in the hydrological model. The quantities delivered by the hydrologic model have to be up-scaled for use in the meteorological model.

To downscale the hydrologically relevant quantities an explicit subgrid-scheme, firstly suggested by Seth et al. (1994) for the global scale, should be adapted for the mesoscale to downscale the hydrologically relevant quantities (e.g., Mölders et al. 1996, 1999, Mölders and Raabe 1997). Herein, a higher resolution grid is defined at the interface earth-atmosphere and within the soil. This higher resolved grid consists of several subgrid cells per grid cell (e.g., Fig. 4). According to the mixture approach suggested before, these subgrid cells may be covered by at least one vegetation- and/or soil-type. Thus, coupled energy (for soil and vegeta-

tion) and hydrological budgets are maintained for each subgrid cell using the subgrid cell surface characteristics and the micro-climate at the representative location. This means that, in each subgrid cell, the fluxes are individually calculated with their own subgrid soil forcing and near-surface meteorological forcing in the immediate vicinity of the earth's surface. Soil water content, soil-temperature, near-surface air-temperature and humidity have to be stored for each subgrid cell. They have to be used to determine these quantities in the next time step. The coupling of the subgrid cells to the atmospheric grid cell may be realized by the arithmetic average of individual subgrid cell fluxes to provide the grid cell fluxes (e.g., Avissar and Pielke 1989, Mölders and Raabe 1996, Mölders et al. 1996).

The explicit subgrid scheme belongs to the class of mosaic approaches. The main advantages of the subgrid scheme as compared to a simple mosaic approach are that by explicitly breaking down the grid cells of the atmospheric model (1) the spatial location of each subgrid flux is known (Fig. 4), (2) precipitation can easily be heterogenized, (3) the coupling can be realized on the subgrid cells, i.e., the hydrological and meteorological model have the same surface characteristics. The disadvantage of the explicit subgrid scheme is that it can be much more computationally expensive than a simple mosaic approach, especially, when the surface conditions are relatively homogeneous (see Mölders et al. 1996).

The precipitation provided by a state-of-the art cloud module by the atmospheric model has to be downscaled (e.g., von Storch et al. 1993, Leung and Ghan 1995) to the resolution of the hydrological model. In most regions of the world, long-lasting precipitation increases with elevation (orographic effect) because horizontally moving air encounters a topographic barrier and, hence, acquires vertical motion when passing the barrier. The related cooling leads to precipitation. Assuming that precipitation increases with elevation and taking into account the direction of wind, stratiform precipitation can be heterogenized (e.g., Leung and Ghan 1995, Leung et al. 1996, Mölders et al. 1996). For convective cases, heterogenization of precipitation should be related to surface characteristics (e.g., large patches of sand) and wind speed and wind direction. Note that, in the meteorological model used for the coupling, a bulk-parameterization (e.g., Mölders et al. 1997) considering at least five water classes (water vapor, cloud water, rainwater, ice, graupel) should be applied to predict the mean precipitation within a meteorological grid cell, because such a parameterization is more physical and closer to the processes than a cumulus-parameterization, for instance (e.g., Mölders et al. 1994).

Short vegetation should be considered by at least one layer. In the case of high vegetation, multi-layer canopy models (e.g., Ziemann 1998) should be applied within the subgrid cells partly or totally covered by high vegetation. Note that, up to now, there exist no NWP models or GCMs which consider high vegetation by some kind of multiple layer canopy model.

Interception loss should be included at least for high vegetation. Herein, it should be allowed that only parts of the vegetation are wetted. In such a case, three energy and two hydrological budget equations have to be solved within a subgrid cell simultaneously. Here, again the mixture approach can be applied to allow partly wetted canopies.

3.3.2 *Soil physics*

In the hydrometeorological module, all soil processes are calculated on the subgrid that ought have the same resolution as the hydrological model. The treatment of the soil physics should be based on the thermodynamics of irreversible processes (e.g., Philip and de Vries 1957, Philip 1957, de Vries 1958). Among other things Darcy's law and the Richards equation should be included. The effect of the horizontal soil water fluxes should be calculated by the hydrological model and can be determined with a coarser time step than that used by the hydrometeorological module and the atmospheric model. The heat- and moisture transport

within the soil should be solved by balance equations for soil-temperature and volumetric water content (e.g., Zdunkowski 1983, Kramm et al. 1996). These equations are a coupled system of partial differential equations and have to be solved iteratively (e.g., Kramm et al. 1996). Solving the coupled equations means that the Ludwig-Soret effect (i.e., a temperature gradient can change soil volumetric water content) and Dufor effect (i.e., a moisture gradient may alter soil temperature) are taken into account. An approximation that decouples moisture and heat processes within the soil should be avoided if long-term simulations were to be carried out that are often the aim of coupled model simulations (e.g., investigations on the impact of land-use changes on climate and water resources). Within the soil, different soil types should be allowed in the vertical soil column. The soil-module should reach into the level of ground water to be able to examine impacts and changes of ground water recharge on soil moisture, evapotranspiration, and runoff.

Moreover, the soil-water uptake by roots should be considered to assure that evapotranspiration is related to the volumetric water content of different soil layers (not only the uppermost soil layer as often realized in meteorological models). Parameterizations using the Cowan (1965) model could be applied (e.g., Martin 1990, Dickinson 1993). It should be ensured that the root amount can vary with depth (e.g., Wilson et al. 1987) and time (e.g., Mölders et al. 1999b).

The parameterization of infiltration should be consistent with Richards equation. It should allow that only parts of the atmospheric grid cell experience infiltration. This may be realized within the frame-work of an explicit subgrid scheme wherein, among others, soil type, land-use type, and precipitation may be heterogenized (e.g., Mölders 1996).

For coupling of a hydrological model with a NWP or GCM, soil frost processes have to be considered, because large parts of the continents are regularly frozen during winter in high and mid-latitudes or in mountainous regions. From the hydrological point of view soil frost leads to the freezing of soil-water for which its mobility is nearly totally restricted and capillarity, infiltration as well as percolation are only slightly effective. Since soil frost hinders infiltration of water into the soil (e.g., Cherkauer and Lettenmaier 1999), rain falling onto frozen soil or melting of a snow package laying over frozen soil will contribute to runoff. Aspects of soil frost that affect the atmosphere are more indirect than those in hydrology. The thermic stability and low air temperatures, and the consequently low saturation pressure of water vapor lead to low evaporation. Thus, the moisture will be stored in frozen soil and may increase spring peak flood events (e.g., Cherkauer and Lettenmaier 1999). In addition, transpiration plays a minor role because deciduous forests have already lost their leaves and even the stomatal conductivity of coniferous forests is low then. Obviously, if the freezing processes of soils are not considered, too high water vapor fluxes into the atmosphere will be predicted as there is seemingly still „liquid“ water available that, moreover, requires less energy for evaporation than ice.

The boundary between an unfrozen upper soil layer and a frozen deeper soil layer, for instance, may vary within the diurnal course. The determination of the surface water and energy fluxes is extremely difficult when the exact depth of the freezing line is unknown. The reduced hydraulic conductivity of frozen soil increases the potential for high snowmelt runoff losses.

For the reasons discussed above, the inclusion of soil frost processes when coupling hydrological and meteorological models is an urgent need for an adequate calculation of runoff and water supply to the atmosphere in winter. The terms of soil frost and thawing should be included in the coupled equations of heat and moisture transport within the soil and should be based on the thermodynamics of irreversible processes. Thus, it has to be run at the same time step than the soil model part of the hydrometeorological model (e.g., Fuchs et al. 1978, Flerchinger and Hanson 1989, Cherkauer and Lettenmaier 1999).

3.3.3 Snow

Another important process to be considered in the hydrometeorological module is the treatment of snowmelt and previous snow accumulation. Snow is commonly treated differently in the hydrological and meteorological models because of the different relevance of the various aspects of snow for meteorological and hydrological processes. In hydrological modeling, namely, the retarded entering of precipitation into the land phase of the water cycle is the most prominent aspect of snow. Thus, a simple day-degree method is sufficient in most hydrological applications. Besides the retarded entering of water in the land-phase of the water cycle, in meteorological models, the insulating effect of snow that prevents the underlying soil from cooling as well as the high albedo of snow that affects the energy budget (e.g., Plüss and Ohmura 1997, Abdalati and Steffen 1997, Cline 1997, Baker et al. 1999, Robinson et al. 1992) are the most important aspects of snow to be considered. Albedo, for instance, dramatically changes when snow falls and rests on the ground, especially, where the underlying ground has albedo below 0.15 when wet. Since the albedo associated with snow cover typically ranges between 0.35 and 0.9, the coupling between the surface and atmosphere is generally weaker than in summer.

Disappearance of snow leads to runoff and removes a critical constraint on both water vapor pressure and surface temperature. As long as the snowpack exists these quantities cannot rise above 610 Pa and 273.15 K. Therefore, the surface-atmosphere coupling should proceed with more vigor after the melting of snow (Baker et al. 1999). Exposed soil surfaces within a partly broken snow coverage lead to substantial sensible heat fluxes, convection, and increased vertical mixing in the surface layer. If sufficient moisture is available, clouds may form. The cloud shadows may feed back to a reduced melt process. The strong spatial contrast in the energy budget of snow-covered and snow-free areas may lead to significant advective flow similar to a sea breeze (Baker et al. 1999).

As a consequence of the aspects discussed above, in coupling a hydrological and a meteorological model, the snow accumulation and melt processes should be considered by a multiple layer snow model (e.g., Anderson 1976, Foster et al. 1996, Cayan 1996) if snow events were frequent and snow accumulation is high. In the case, that the coupled model is mainly applied in regions of seldom snowfall and usually not long-lasting snow coverage it has to be examined whether a single layer model can be sufficient.

4. Conclusions

During the last decade, several attempts to consider the interaction between the land and atmospheric part of the water cycle were undertaken in both long-term climate modeling and short-term weather forecast applications. There are three different concepts of how to treat the complexity of the physical system “water cycle”:

- parameterization of subsurface and surface hydrological processes in the atmospheric model,
- one-way or two-way coupling of hydrological and atmospheric models by data exchange, and
- direct coupling of hydrological and atmospheric models by use of a common intersection, here denoted as hydrometeorological module.

The latter seems to be the most advantageous way is to realize the coupling between hydrological and meteorological models, because such a realization would save computer capacities as well as man-power and provides an optimized physical consistency between the models.

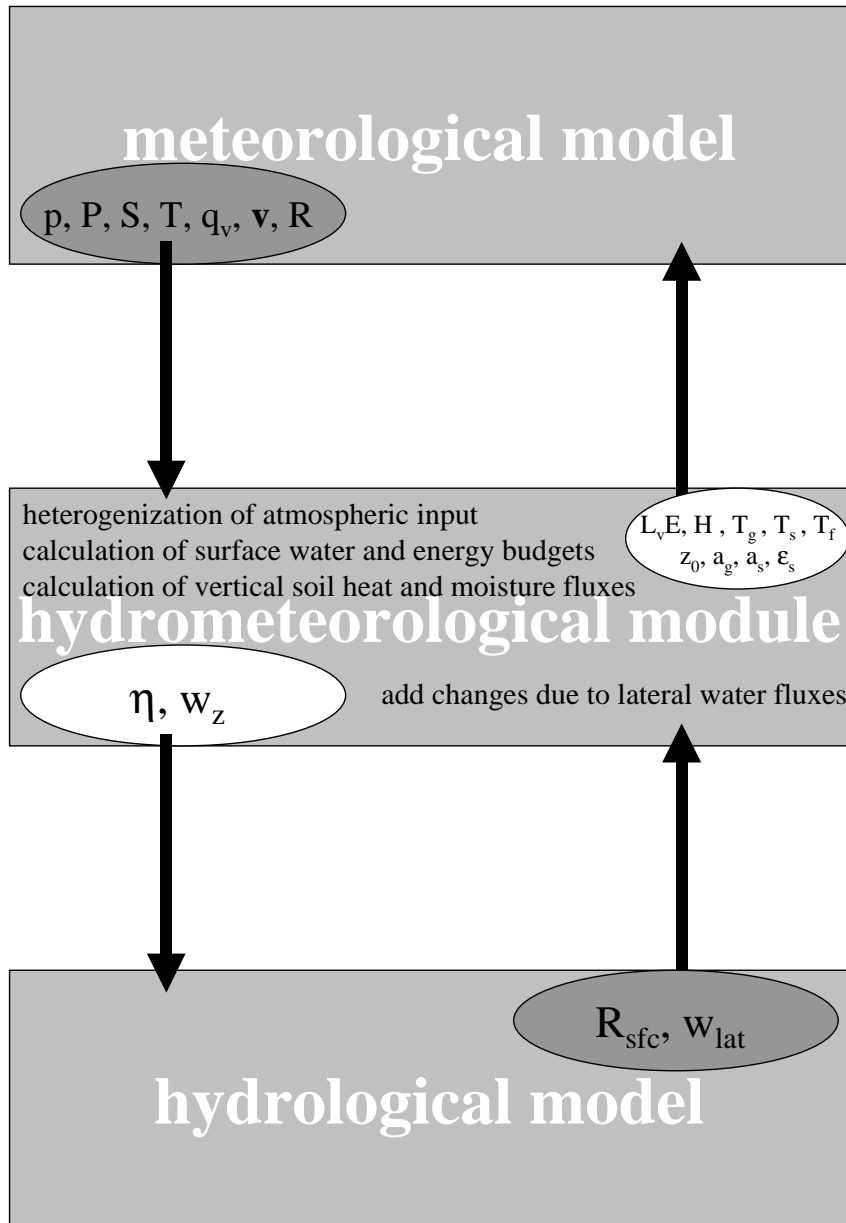


Fig. 7. Flow chart of calculated data exchange in a hydrometeorological module used as intersection for coupling a hydrological and meteorological model. Here, $p, S, T, T_g, T_s, T_f, q_v, \mathbf{v}, R, \epsilon_s, \alpha_g, \alpha_s, \eta, w_{lat}, w_z, R_{sfc}, L_v E, H, T_g, T_s, T_f, z_0, \alpha_g, \alpha_s, \epsilon_s$ stand for pressure, precipitation (rain and/or snow), the temperatures of the air, ground, snow and foliage, specific humidity of air, wind vector, short-wave and counter radiation radiation, emissivity of snow, albedo of ground and snow, volumetric water content, horizontal and vertical moisture and water fluxes in the soil, (channel and/or surface) runoff, latent and sensible heat flux densities, ground temperature, snow temperature foliage temperature, roughness length, albedo of the ground and snow as well as emissivity of snow, respectively.

A concept of the design of a hydrometeorological module was introduced. Such a hydrometeorological module is to serve as the lower boundary condition for the atmospheric model and as upper boundary condition for the hydrological model. The hydrometeorological module is to be called for each surface grid cell of the atmospheric model and should include subgrid cell representation of prognostic snow-cover, prognostic equation for soil volumetric

water content and soil temperatures (in z-direction only) under consideration of the Ludwig-Soret- and Dufor-effect, treatment of soil freezing and thawing, water uptake by roots, local runoff of heavy precipitation and snowmelt, and of energy and moisture budgets for soil, vegetation, canopy air, temporary surface water (e.g., intercepted water, flood, snow-cover). These subgrid cells should match the resolution of the hydrological model. Soil and snow-cover are to be divided into multiple vertical levels. Vegetation and canopy are to be represented by at least a single layer.

The hydrometeorological module provides the water and energy fluxes, surface temperature, and moisture, surface albedo and emissivity to the atmospheric model, while it provides the vertical soil water fluxes, soil volumetric water content, ground water recharge, infiltration, melt water, and ponded water to the hydrological (Fig. 7). Note that the ponded water can produce runoff. The atmospheric model delivers to the hydrometeorological module surface pressure, specific humidity, air temperature, wind, and short- and long-wave downward radiation (Fig. 7). The hydrological model provides to the hydrometeorological module the lateral soil water fluxes, surface and channel runoff (Fig. 7).

Acknowledgements

I would like to express my thanks to DFG for financial support of this concept under contract Mo770/2-1. I wish to thank A. Ziemann for fruitful discussion and helpful comments.

References

- Abbott, M.B., J.C. Bathurst, J.A. Cunge, P.E. O'Connell, J. Rasmussen, 1986. An introduction to the european hydrological system - Systéme Hydrologique Européen, „SHE“, 2: structure of a physically-based, distributed modelling system. *J. Hydrol.* **87**, 61-77.
- Abdalati, W., and Steffen, K., 1997. Snowmelt on the Greenland ice sheet as derived from passive microwave satellite data. *J. Clim.* **10**, 165-175.
- Anderson, E.A., 1976. *A point energy and mass balance model of snow cover*. NOAA Technical Memorandum NWS Hydro-17. US Department of Commerce, Silver Springs, MD, pp. 217.
- Avissar, R. and Pielke, R.A., 1989. A parameterization of heterogeneous land surface for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.* **117**, 2113-2136.
- Baker, J.M., K.J. Davis, and G.C. Liknes, 1999. Surface energy balance and boundary layer development during snowmelt. *J. Geophys. Res.* **104D**, 19611-19621.
- Band, L.E., Patterson, P., Nemani, R., and Running, S.W., 1993. Forest ecosystem processes at the watershed scale: Incorporating hillslope hydrology. *Agric. For. Meteor.* **63**, 93-126.
- Beven, K.J., and Kirby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* **24**, 43-69.
- Cayan, D., 1996. Interannual climate variability and snowpack in the western United States. *J. Climate* **9**, 928-948.
- Chen, F., K. Mitchell, J. Schaake, Y. Xue, H.-L. Pan, V. Koren, Q.Y. Duan, M. Ek, and A. Betts, 1996. Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.*, **101D**, 7251-7268.
- Cherkauer, K.A., and Lettenmaier, D.P., 1999. Hydrologic effects of frozen soils in the upper Mississippi river basin. *J. Geophys. Res.* **104D**, 19611-19621.
- Cline, D.W., 1997. Effect of seasonality of snow accumulation and melt on snow surface energy exchanges at a continental alpine site. *J. Appl. Meteor.* **36**, 32-51.
- Cowan, I.R., 1965. Transport of water in soil-plant-atmosphere system. *J. Appl. Ecol.* **2**, 221-239.
- Deardorff, J.W., 1978. Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. *J. Geophys. Res.* **84C**, 1889-1903.
- de Vries, D.A., 1958. Simultaneous transfer of heat and moisture in porous media. *Transactions Am. Geophys. Union* **39**, 909-916.
- Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J. and Wilson, M.F., 1986. Biosphere-atmosphere transfer scheme (BATS) for the NCAR community climate model. *NCAR Technical Note 275+STR*.
- Dickinson, R.E., 1993.

- Dolman, A., 1992. A note on areally-averaged evaporation and the value of the effective surface conductance. *J. Hydrol.* **138**, 583-589.
- Entekhabi, D. and Eagleson, P., 1989. Land surface hydrology parameterization for atmospheric general circulation models including subgrid-scale spatial variability. *J. Clim.* **2**, 816-831.
- Famiglietti, J., and Wood, E.F., 1991. Evapotranspiration and runoff for large land areas: land surface hydrology for atmospheric general circulation models. *Surv. Geophys.* **12**, 179-204.
- Flerchinger, G.N., and Hanson, C.L., 1989. Modeling soil freezing and thawing on a rangeland watershed. *Transactions ASAE* **32**, 1551-1554.
- Foster, J., Liston, G., Koster, R., Essery, R., Behr, H., Dümenil, L., Verseghy, D., Thompson, S., Pollard, D., and Cohen, J., 1996. Snow cover mass intercomparison of General Circulation Models and remotely sensed datasets. *J. Climate* **9**, 409-426.
- Friedrich, K., Mölders, N., and Tetzlaff, G., 2000. On the influence of surface heterogeneity on the Bowen-ratio: A theoretical case study. *Theor. Appl. Clim.* **65**, 181-196.
- Fuchs, M., Campbell, G.S., and Papendick, R.I., 1978. An analysis of sensible and latent heat flow in partially frozen unsaturated soil. *Soil Sci. Soc. Am. J.* **42**, 379-385.
- Giorgi, F., R. Avissar, 1997. Representation of heterogeneity effects in earth system modeling: Experience from land surface modeling. *Rev. Geophys.* **35**, 413-438.
- Graham, L.P., and S. Bergström, 2000. Land surface modelling in hydrology and meteorology – lessons learned from the Baltic Basin. *Hydrology and Earth System Sciences*, **4**, 13-22.
- Hagemann, S., Dümenil, L., 1998. A parameterization of the lateral waterflow for the global scale. *Clim. Dyn.* **14**, 17-31.
- Kramm, G., Dlugi, R., Mölders, N., and Müller, H., 1994. Numerical investigations of the dry deposition of reactive trace gases. In: J.M. Baldasano et al. (eds.) *Computer Simulation Air Pollution II* Vol. 1, Computational Mechanics Publications, Southampton Boston, 285-307.
- Kramm, G., Beier, N., Foken, T., Müller, H., Schröder, P., and Seiler, W., 1996. A SVAT scheme for NO, NO₂, and O₃ - Model description and test results. *Meteorol. Atmos. Phys.* **61**, 89-106.
- Kuhl, S.C., and Miller, J.R., 1992. Seasonal river runoff calculated from a global atmospheric model. *Water Resources Res.* **28**: 2029-2039.
- Leung, L.R., and Ghan, S.J., 1995. A subgrid parameterization of orographic precipitation. *Theor. Appl. Climatol.* **52**, 95-118.
- Leung, L.R., Wigmosta, M.S., Ghan, S.J., Epstein, D.J., and Vail, L.W. 1996. Application of a subgrid orographic precipitation/surface hydrology scheme to a mountain watershed. *Geophys. Res.*
- Lhomme, J.-P., 1992. Energy balance of heterogeneous terrain: averaging the controlling parameters. *Agri. For. Meteorol.* **61**, 11-21.
- Liston, G.E., Sud, Y.C. and Wood, E.F., 1994. Evaluating GCM land surface hydrology parameterizations by computing river discharges using a runoff routing model: Application to the Mississippi Basin. *J. Appl. Met.* **33**, 394-405.
- Marengo, J.A., Miller, J.R., Russel, G.L., Rosenzweig, C.E. and Abramopoulos, F., 1994. Calculations of river-runoff in the GISS GCM: impact of a new land-surface parameterization and runoff routing model on the hydrology of the Amazon river. *Clim. Dyn.* **10**, 349-361.
- Martin, P. 1990. Forst succession and climatic change: coupling land-surface processes and ecological dynamics. Ph.D. Thesis *NCAR/CT-131*.
- Mengelkamp, H.-T., Warrach, K., and Raschke, E., 1999. SEWAB – a parameterization of the surface energy and water balance for atmospheric and hydrologic models. *Advances in Water resources* **23**, 165-175.
- Miller, J.R., Russel, G.L., and Caliri, G., 1994. Continental-scale river flow in climate models. *J. Clim.* **7**, 914-928.
- Milly, P.C.D., and Dunne, K.A., 1994. Sensitivity of the global water cycle to the water-holding capacity of land. *J. Clim.* **7**: 506-526.
- Mölders, N., Hass, H., Jakobs, H.J., Laube, M., and Ebel, A., 1994. Some effects of different cloud parameterizations in a mesoscale model and a chemistry transport model. *J. Appl. Meteor.* **33**, 527-545.
- Mölders, N., and Raabe, A., 1996. Numerical investigations on the influence of subgrid-scale surface heterogeneity on evapotranspiration and cloud processes. *J. Appl. Meteor.* **35**, 782-795.
- Mölders, N., Raabe, A., and Tetzlaff, G., 1996. A comparison of two strategies on land surface heterogeneity used in a mesoscale β meteorological model. *Tellus* **48A**, 733-749.
- Mölders, N., Kramm, G., Laube, M., and Raabe, A., 1997. On the influence of bulk-parameterization schemes of cloud microphysics on the predicted water-cycle relevant quantities - a case study. *Met. Zeitschr.* **6**, 21-32.
- Mölders, N., and Raabe, A., 1997. Testing the effect of a two-way-coupling of a meteorological and a hydrologic model on the predicted local weather. *Atmos. Res.* **45**, 81-108.

- Mölders, N., 1999. On the effects of different flooding stages of the Odra and different landuse types on the local distributions of evapotranspiration, cloudiness and rainfall in the Brandenburg-Polish border area. *Contrib. Atmos. Phys.* **72**, 1-24.
- Mölders, N., Raabe, A., and Beckmann, T., 1999a. A technique to downscale meteorological quantities for the use in hydrologic models - Description and first results. *IAHS* **254**, 89-98.
- Mölders, N., Haferkorn, U., Knappe, S., Döring, J., Kramm, G., 1999b. Evaluation of simulated water budget by means of measurements at Brandis lysimeter station. In: Tetzlaff, G., U. Grünwald (Eds.), 2. Tagung des Fachausschusses Hydrometeorologie am 15./16. November 1999 in Leipzig, Wiss. Mitt. Leipzig, 16, 67-83.
- Mölders, N., and Rühaak, W., 2001. Sensitivity studies with a surface and channel runoff module coupled to a mesoscale atmospheric model. In: Arnold, K. (Ed.), Wissenschaftliche Arbeiten aus Leipzig, Mitt. Inst. Meteor., 23, pp. 12.
- Müller, E., Foken, T., Heise, E. and Majewski, D., 1995. LITFASS - A nucleus for a BALTEX field experiment. *DWD Arbeitsergebnisse* No. 33.
- Philip, J.R., and de Vries, D.A., 1957. Moisture in porous materials under temperature gradients. *Transactions Am. Geophys. Soc.* **18**, 222-232.
- Philip, J.R., 1957. Evaporation, and moisture and heat fields in the soil. *J. Met.* **14**, 354-366.
- Plüss, C., and A. Ohmura, 1997. Longwave radiation on snow-covered mountainous surfaces. *J. Appl. Meteor.* **36**, 818-824.
- Refsgaard, J.C., Seth, S.M., Bathurst, J.C., Ehrlich, M., Storm, B., Jørgensen, G.H., and Chandra, S., 1992. Application of SHE to catchment in India - Part 1: General results. *J. Hydrol.* **140**, 1-23.
- Robinson, D.A., Serreze, M.C., Barry, R.G., Scharfen, G., and Kukla, G., 1992. Large scale patterns and variability of snowmelt and parameterized surface albedo in the Arctic Basin. *J. Climate* **5**, 1109-1119.
- Sausen, R., Schubert, S., and Dümenil, L., 1994. A model of the river runoff for use in coupled atmosphere-ocean models. *J. Hydrol.* **155**, 337-352.
- Sellers, P.J., Mintz, Y., Sud, Y.C. and Dalcher, A., 1986. A simple biosphere atmosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.* **43**, 505-531.
- Seth, A., Giorgi, F., and Dickinson, R.E., 1994. Simulating fluxes from heterogeneous land surfaces: explicit subgrid method employing the biosphere-atmosphere transfer scheme (BATS). *J. Geophys. Res.* **99D**, 18651-18667.
- Shuttleworth, W.J., 1988. Macrohydrology - the new challenge for process hydrology. *J. Hydrol.* **10**, 31-56.
- Stieglitz, M., Rind, D., Famiglietti, J.S., and Rosenzweig, C., 1997. An efficient approach to modeling the topographic control of surface hydrology for regional and global climate modeling. *J. Climate* **10**, 118-137.
- Todini, E., 1988. Rainfall-runoff modelling - past, present and future. *J. Hydrol.* **100**, 341-352.
- von Storch, H., Zorita, E., and Cubasch, U., 1993. Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *J. Clim.* **6**, 1161-1171.
- Walko, R.L., Band, L.E., Baron, J., Kittel, T.G.F., Lammers, R., Lee, T.J., Ojima, D., Pielke, R.A. Sr., Taylor, C., Tague, C., Tremback, C.J., and Vidale, P.L., 2000. Coupled atmosphere-biophysics-hydrology models for environmental modeling. *J. Appl. Meteor.* **39**, 931-944.
- Wetzel, P.J., and Chang, J.-T., 1988. Evapotranspiration from nonuniform surfaces: A first approach for short-term numerical weather prediction. *Mon. Wea. Rev.* **116**, 600-621.
- Wilson, M.F., Henderson-Sellers, A., Dickinson, R.E., and Kennedy, P.J., 1987. Sensitivity of the biosphere-atmosphere transfer scheme (BATS) to the inclusion of variable soil characteristics. *J. Clim. Appl. Met.* **26**, 341-362.
- Zdunkowski, W., 1983. Transport equations for heat and moisture in the soil and their application to boundary layer problems. *Contr. Phys. Atmos.* **56**, 58-83.
- Ziemann, A., 1998. Numerical simulations of meteorological quantities in and above forest canopies. *Meteorol. Zeitsch.* **7**, 120-128.

Author's address:

LIM - Institut für Meteorologie, Universität Leipzig, Stephanstraße 3, 04103 Leipzig

Sensitivity studies with a surface and channel runoff module coupled to a mesoscale atmospheric model

N. Mölders and W. Rühaak

Summary:

A module to investigate ground water recharge was developed, and implemented into the mesoscale meteorological model GESIMA (Geesthacht's simulation model of the atmosphere) as well as coupled to a soil-vegetation scheme. Important features of the ground water module are the determination of surface and channel runoff. A comparison of the results provided by GESIMA with and without consideration of surface and channel runoff shows a remarkable impact of surface runoff on the soil moisture fluxes. Substituting water meadows by willow-forests demonstrates their importance for soil moisture fluxes.

Zusammenfassung:

Ein Modul zur Untersuchung von Grundwasserneubildung wurde entwickelt, in das mesoskalige meteorologische Modell GESIMA (Geesthachter Simulationsmodell der Atmosphäre) integriert und an ein Boden-Vegetationsmodell gekoppelt. Wesentliche Bestandteile des Grundwassermoduls sind die Berechnung des Oberflächen- und Gerinneabflusses. Ein Vergleich der Ergebnisse von GESIMA, die mit und ohne Oberflächen- und Gerinneabfluss erstellt wurden, belegt einen deutlichen Einfluss des Oberflächenabflusses auf die berechneten Feuchteflüsse im Boden. Untersuchungen zum Einfluss von Auenwäldern auf die Grundwasserneubildung belegen deren Bedeutung für die Wasserflüsse im Boden.

1. Introduction

In the last 70 years, the human population increased three times. Under an increasing population as well as changing climate and land-use conditions the availability of blue water (ground water) and green water (plant available water for food production) becomes a central question of research interest. Water availability is not only affected by climate conditions, but also by the surface conditions and, hence, land-use. It, among other things, depends on precipitation, infiltration, surface runoff, as well as on the exchange of heat, matter, and momentum at the earth-atmosphere interface. Due to the dependence of water availability on both the land-part and the atmospheric part of the water cycle studies on water availability require a closed consideration of the water cycle. The neglecting of lateral soil water movements and surface runoff, for instance, may yield to an underestimation of soil moisture in river valleys and an overestimation of soil moisture in the nearby mountainous regions which usually receive more precipitation (e.g., Müller et al. 1995). A wrong soil moisture distribution affects the water supply to the atmosphere and, hence, may appreciably affect the local recycling of precipitation (Mölders 1999).

In atmospheric models, however, surface runoff, the transport of water by river flow as well as the re- and discharge of groundwater are commonly neglected. Recently, to more appropriately model the water cycle, several authors (e.g., Kuhl and Miller 1992, Marengo et al. 1994, Miller et al. 1994, Sausen et al. 1994, Hagemann and Dümenil 1998) suggested parameterizations of different complexity to directly parameterize runoff in general circulation models (GCM). Another way to model more appropriately the water cycle is to directly couple a

hydrologic and a meteorological model in a two-way manner exchanging data of precipitation, soil moisture, evapotranspiration, lateral interflow, and surface runoff (e.g., Mölders and Raabe 1997, Mölders et al. 1999a). Herein, the hydrologic processes of the river catchment are considered in the atmospheric model, which itself drives the hydrologic model. First results of such a loosely coupled hydrologic-meteorological model package showed that, especially, the contribution of short, but intensive showers to runoff will be underestimated, if the hydrological and meteorological model only exchange data hourly (Mölders and Raabe 1997, Mölders et al. 1999a). Therefore, even though some hydrologic processes (e.g., lateral soil water fluxes, groundwater re- or discharge) could be determined with much larger time steps, these processes as well as surface and channel runoff should be included into the soil vegetation atmosphere transfer scheme applied in the atmospheric model. Such a concept is realized and tested in our study.

To examine the role of surface runoff in the regional water cycle and its meaning for water availability under changed land-use conditions a module to describe surface and channel runoff as well as ground water recharge within a mesoscale atmospheric model was developed and coupled to a state-of-the-art hydro-dynamics soil-vegetation scheme.

2. Brief description of the module for surface and channel runoff

The temporal and spatial scales required for surface and channel runoff on one hand side and the atmospheric model on the other differ. Unfortunately, in meteorological models a finer grid resolution is restricted by parameterization limitations (e.g., the fetch-conditions from which the parameterizations of the surface fluxes are derived lead to a requirement of a relation of 1:100 for the ratio of the vertical to horizontal grid resolution), the limited availability of initial data, and insufficient computer resources. In order to couple the runoff module to GESIMA the scale differences have to be bridged by aggregation/disaggregation. Herein, a higher resolution grid is defined at the interface earth-atmosphere and within the soil. This higher resolved grid consists of several subgrid cells per grid cell (see Mölders et al. 1996, 1999a).

Precipitation rate is one water cycle relevant input quantity to be heterogenized. The cloud module provides this quantity on the resolution of the atmospheric model which amounts to $5 \times 5 \text{ km}^2$ in our study. In most regions of the world, long-lasting precipitation increases with elevation (orographic effect) because horizontally moving air acquires vertical motion when it encounters and passes a topographic barrier. The related cooling may lead to precipitation. On average, in Mid-Europe, the annual precipitation increases 50 mm per 100 m terrain elevation. This gradient is stronger for maritime than for continental conditions (Pleiss 1977). According to these findings precipitation is heterogenized by relating precipitation to subgrid-scale surface elevation (Mölders et al. 1996)

$$P_i = P_t \cdot \left(\frac{z_i}{z} \right) \quad (1)$$

where

$$\frac{-}{z} = \frac{1}{N} \cdot \sum_{i=1}^N z_i \quad (2)$$

and z_i are the mean terrain height of the respective grid cell and the i th subgrid cell, P_t is the mean precipitation predicted by the cloud module for the respective grid cell, and N stands for the number of subgrid cells (= 25 in this study). The relationship between the ponded water on the coarse grid is realized to that on the fine grid in a similar manner as for precipitation. Herein, of course, it is ensured that a subgrid cell that has no ponded water at time $t-1$ will not

have water at time t except from precipitation in this subgrid cell or by receiving a surface water flow from one of its higher elevated neighbored subgrid cells.

Besides all the representative characteristics of a subgrid cell (like elevation, land-use type, soil type, etc.) its simulated quantities (e.g., precipitation, infiltration, ponded water, runoff) have to be stored for each subgrid cell. These values have to be used to determine these quantities in the next time step. The coupling of the subgrid cells to the atmospheric grid cell is realized by the arithmetic average of individual subgrid cell fluxes and variables of state to provide the grid cell fluxes and variables of state (see Mölders et al. 1996, 1999a, Mölders and Rühaak 2000).

Surface runoff describes the two-dimensional water flow on a sloped surface. By assuming that friction is dominant surface runoff can be calculated by the St.Vernant equation which is a partial quasi-linear differential equation of hyperbolic type (Maniak 1993). By considering the sink and source terms it can be written as (e.g., Lawler 1964, Abbott et al. 1986)

$$\frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} - q = 0 \quad . \quad (3)$$

Herein, $h(x,y)$ is the local water depth, t is time, $u(x,y)$ and $v(x,y)$ are speed of flow in x - and y -direction, $q(x,y,t)$ is the effective precipitation minus infiltration, I_g . By applying the Strickler-Manning-resistance law one obtains a relation between the velocity and depth of water (e.g. Abbott et al. 1986)

$$uh = k_x I_x^{1/2} h^{5/3} \quad , \quad (4)$$

$$vh = k_y I_y^{1/2} h^{5/3} \quad , \quad (5)$$

where $k_x(x,y)$ and $k_y(x,y)$, $I_x(x,y)$ and $I_y(x,y)$ stand for the Manning-Strickler-coefficients and the gradients of the water surface in the x - and y -directions, respectively. In the following subscripts, x and y , are dropped when used in general or the direction is clear. Equations 4 and 5 are inserted into equation 3 and are solved by an explicit finite-differences scheme. The Manning-Strickler-coefficients applied in this study are given in table 1.

To determine channel flow the one-dimensional continuity equation (here exemplarily given for the x -direction)

$$\frac{\partial(uA)}{\partial x} + \frac{\partial A}{\partial t} = 0 \quad , \quad (6)$$

is combined with the one-dimensional form of the St.Vernant equation leading to

$$\frac{1}{g} \cdot \frac{\partial u}{\partial t} + \frac{u}{g} \frac{\partial u}{\partial x} + \frac{\partial h}{\partial x} - (I_s - I_r) + \left(\frac{u}{g \cdot A} \cdot q \right) = 0 \quad . \quad (7)$$

Here, A is the area of the channel cross-section and u is the velocity of the channel flow, g is the acceleration of gravity, I_r and I_s are calculated using the Manning-Strickler-coefficients. The technical-numerical approach of equation 7 follows Graf (1998). Trapezoid, rectangular, and triangular channel types can be distinguished. The geometric relation of a trapezoid-formed channel can be formulated by the area of the cross-section, $A = (b + m \cdot h) \cdot h = b \cdot h + m \cdot h^2$, and the width of the water surface, $B = b + 2 \cdot m \cdot h$. The hydraulic depth, D_h , is given by

$$D_h = \frac{A}{B} = \left(\frac{(b + m \cdot h)h}{b + 2 \cdot m \cdot h} \right) \quad . \quad (8)$$

Note that, in the case of a rectangular channel, m is equal to zero and the hydraulic radius equals the hydraulic depth $D_h = h$. In the case of a triangular channel, m is greater than zero

and b is equal to zero. Consequently, the hydraulic depth corresponds to half of the flow depth $D_h = h/2$. Again an explicit finite-differences scheme is used. In doing so, the spatial derivations are determined by weighted centered differences.

The river slopes are estimated by the topographic slopes. A calibration using the known values of river flow, $Q = u \cdot A$, and h can be performed by using the Manning-Stickler-flow formula (e.g., Schröder 1994). Evaporation from the water surface of the river is neglected. Note that, in the subgrid cells, the water surface of the rivers considered in this study are small as compared to the land-surface of the subgrid cells.

The river network is divided into pieces on different flow direction, length, slope, and boundaries conditions for inflow and outflow. Each subgrid cell has its own in- and outflow conditions. Moreover, at the boundaries of the model domain, in- and outflow have to be prescribed by measured values or reasonable values. For purpose of simplicity and easy evaluation of the model results a constant water flux into the rivers that flow into the model domain is assumed at its boundaries. In principal, inflow into a subgrid cell can take place directly from outside of the model domain (e.g., the river Spree), from one of its eight neighbored subgrid cells or by two river joining into the subgrid cell under consideration. River outflow can occur directly out of the model domain or out of the subgrid cell into one of its eight neighbors. In the latter case, the outflow may join into another river (e.g., Schwarze Elster and Elbe). At the boundary of the model domain, river outflow depends on the flow conditions within the last subgrid cell of the channel part, and, hence, may vary with time. The rivers represent sinks for surface runoff. Since no flood event is examined in our study, rivers do not serve as sources for surface runoff.

Tab. 1 Manning-Stickler-coefficients as used in this study. Note that the Stickler-coefficient k is related to Manning's n by $k = 1/n$. The letter d represents the Julian day. Parameters are taken from Maniak (1993), Dingman (1993) and literature cited therein.

Land-use type	Strickler's k ($m^{1/3}/s$)
River	25.00-33.33
Sea, lake	72.46
Open-pit mine, sand, marsh	90.58
Grassland ($z_0 < 0.02m$)	60.39
Grassland ($z_0 \geq 0.02m$)	51.76
Agriculture	51.76
Agriculture in winter ($214 < d < 11$)	60.39
Agriculture in spring ($86 < d < 147$)	45.29
Heath, bushland	30.19
Heath in winter ($214 < d < 11$)	36.23
Deciduous forest	12.08
Mixed forest	15.10
Coniferous forest	18.12
Village	60.39
City	51.76

3. Design of the study

The module to determine surface and channel runoff is introduced into the mesoscale meteorological model GESIMA (Geesthacht's simulation model of the atmosphere). The main dynamical, numerical and physical features of GESIMA are given in Kapitza and Eppel (1992) and Eppel et al. (1995). However, the model physics used in this study differs from that described by Eppel et al. (1995) by (1) the inclusion of an explicit subgrid-scheme to consider subgrid-scale heterogeneity of precipitation and surface as well as channel runoff in a finer

resolution than the grid resolution of the atmospheric model (e.g., Mölders et al. 1996, Mölders and Rühaak 2000); (2) the determination of the surface stress and near-surface fluxes of heat and water-vapor according to Kramm (1989); (3) the treatment of cloud microphysics by a five-water-class bulk-parameterization scheme which includes water-vapor, cloud water, rainwater, ice and graupel (e.g., Mölders et al. 1997); and (4) the soil-vegetation-atmosphere transfers scheme, HTSVS (hydro-thermodynamic soil vegetation scheme; Kramm et al. 1994, 1996) which was included and further-developed by introducing, among other things, a sophisticated parameterization of infiltration (see Mölders 1999, Mölders et al. 1999b, 2000). More details of the model configuration as applied in this study can be found in Mölders (1999).

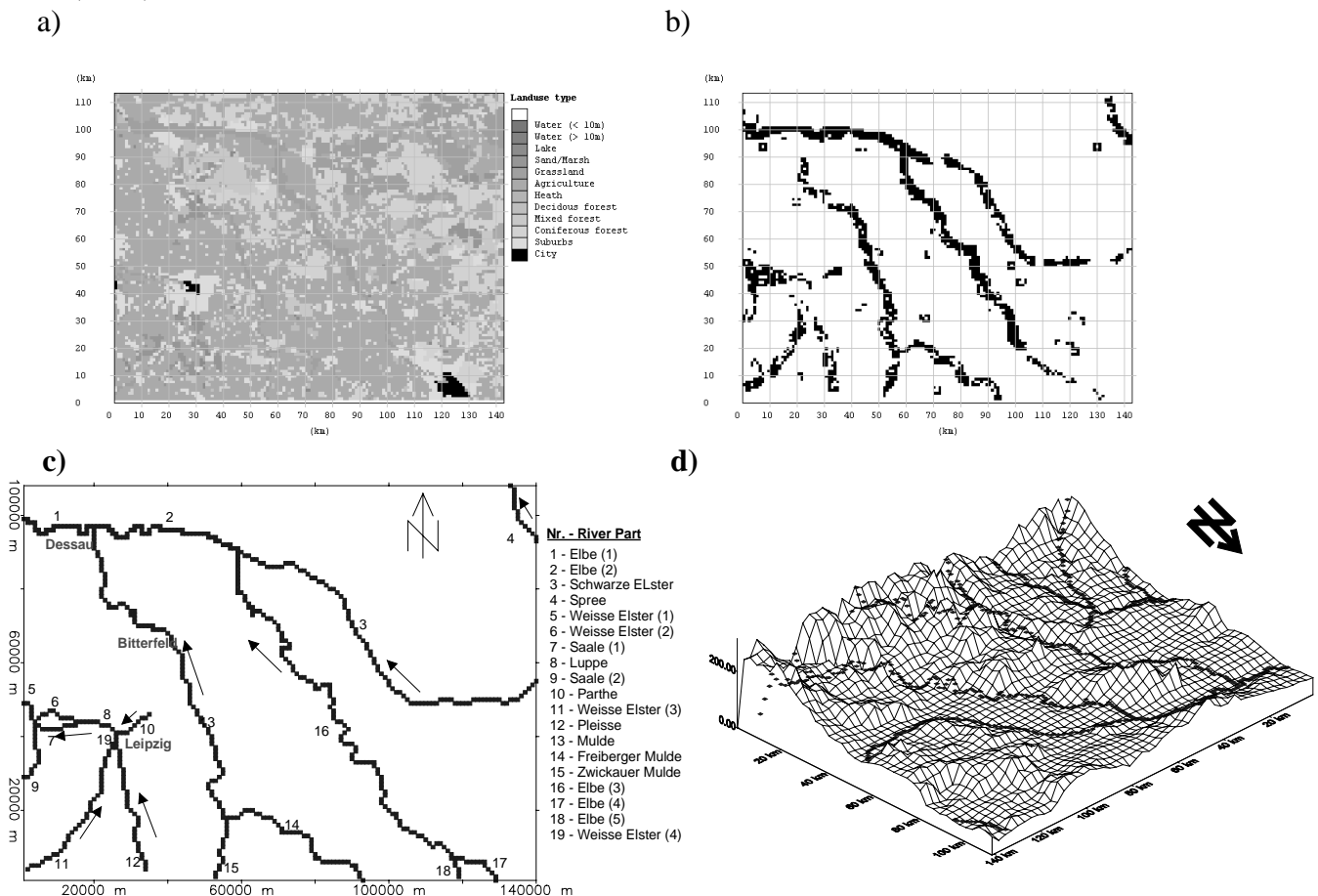


Fig. 1 Schematic view of (a) land-use distribution, (b) location of land-use changes in favor of willow-forest, (c) river network, and (d) terrain elevation as considered in the simulations.

The coupling of the surface and channel runoff modules is realized within the framework of HTSVS. In HTSVS, among other things, the Richards equation is used. The heat- and moisture-processes occurring within the soil are solved by balance equations for soil-temperature and volumetric water content that represent a coupled system of partial differential equations which have to be solved iteratively (e.g., Kramm et al. 1996). The Ludwig-Soret- and Dufor-effect are considered. The parameterization of vegetation includes a mixture approach to consider simultaneously bare soil and/or vegetation within the grid-cell, i.e., the heterogeneity on the microscale is taken into account, which is of relevance for the near-surface stratification of the atmosphere (stability) and the atmospheric fluxes of sensible and latent heat.

GESIMA is run with a horizontal grid resolution of $5 \times 5 \text{ km}^2$. Surface and channel runoff are simulated on a subgrid having a horizontal resolution of $1 \times 1 \text{ km}^2$. Simulations with and without surface and channel runoff are performed for an area of southern Brandenburg and

northern Saxony (Fig. 1). These simulations and their results are referred to as *ho* and *so*, hereafter (Tab. 2). In addition, these simulations are performed for a modified landscape wherein the water meadows along the rivers are changed to willow-forest (Fig. 1), denoted as *hm* and *sm*, respectively (Tab. 2). By assuming that there are no sinks below 1 m depth soil water fluxes at that depth can be regarded as ground water recharge (Haferkorn 1999; personal communication) and are used to evaluate the impact of surface and channel runoff as well as of willow-forests on ground water production.

At the top of GESIMA, a rigid lid together with a sponge layer to absorb vertically propagating gravity waves, are used. At the lateral boundaries, the Orlanski (1976) radiation-boundary condition is applied for the normal component of momentum, and a zero-gradient method is used for all other variables.

The simulated synoptic situation was governed by an anticyclone over the Atlantic, an anticyclone over Russia, and a low pressure system over Germany that hardly moved northwards and led to long-lasting precipitation in the domain of interest.

Tab. 2. Names of simulations and their results.

Simulation	abbreviation
with surface and channel runoff, no land-use changes	ho
with surface and channel runoff, with land-use changes	hm
without surface and channel runoff, no land-use changes	so
without surface and channel runoff, with land-use changes	sm

4. Results and discussion

Primary differences occur due to runoff (e.g., Fig. 2). Whenever or wherever precipitation rate exceeds infiltration rate, water is ponded on the surface and will contribute to runoff in *ho* or *hm*. Neglecting surface runoff means that the water will be ponded if precipitation rate exceeds infiltration rate, and the ponded water will be infiltrated later in *so* or *sm*. Consequently, in the areas of strong precipitation, the total infiltrated water amount of *so* and *sm* exceeds that of *ho* or *hm*, respectively. On the contrary, in the latter simulations, infiltration also occurs in regions without precipitation that are downhill of the area receiving precipitation. Furthermore, since water flows downhill, it can reach areas where precipitation rate is less than possible infiltration rate. Thus, here infiltration is higher in the simulations with consideration of runoff than in those without (e.g., Fig. 3).

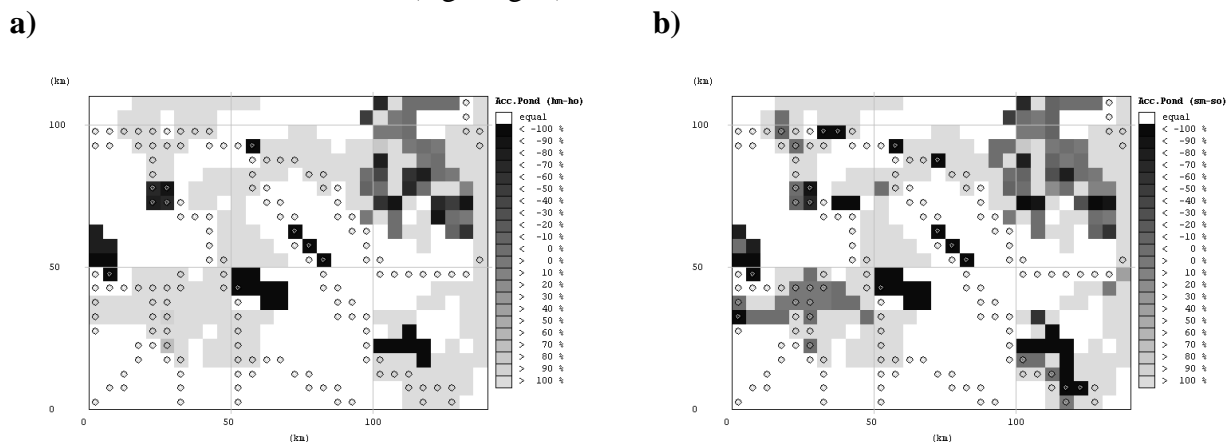


Fig. 2 Percentage difference of accumulated ponded water on the coarse resolution of the atmospheric model after 24 hours of integration as obtained by (a) *hm* – *ho* scaled by *ho*, and (b) *sm* – *so* scaled by *so*. Note that the ponded water can only flow laterally in the simulations with the surface and channel runoff module, i.e. in *hm* and *ho*, respectively. Circles mark grid cells with rivers.

The altered surface moisture leads to secondary differences in water availability, because of the altered evapotranspiration (e.g., Fig. 4). Thus, different amounts of water vapor are supplied to the atmosphere and modify cloud- and precipitation-formation. This means that the simulated local recycling of previous precipitation is affected by the inclusion or neglecting of surface runoff. The altered precipitation (Fig. 5) again yields to differences in water availability (e.g., Fig. 6).

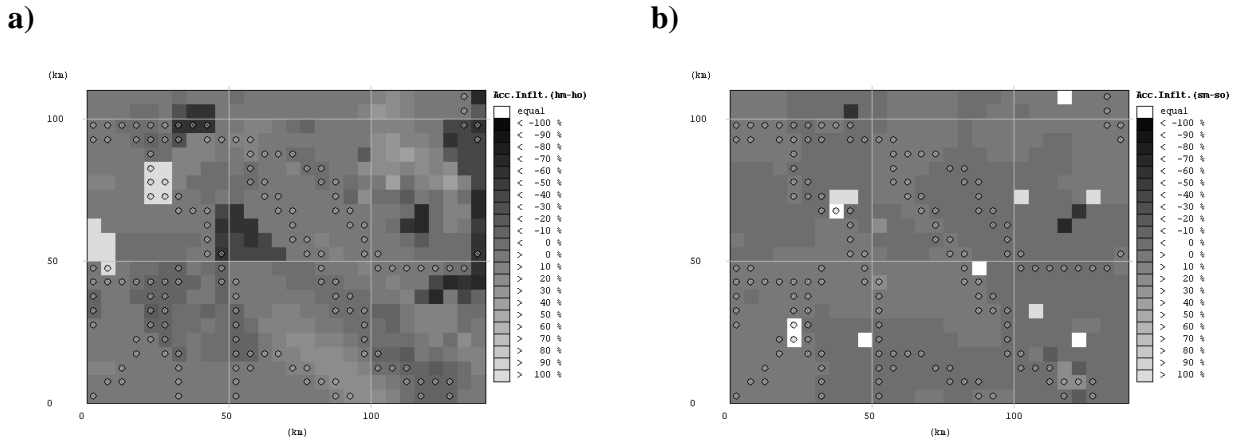


Fig. 3 Like Fig. 2, but for percentage difference of accumulated infiltration after 24 hours of integration as obtained by (a) $hm - ho$ scaled by ho , and (b) $sm - so$ scaled by so .

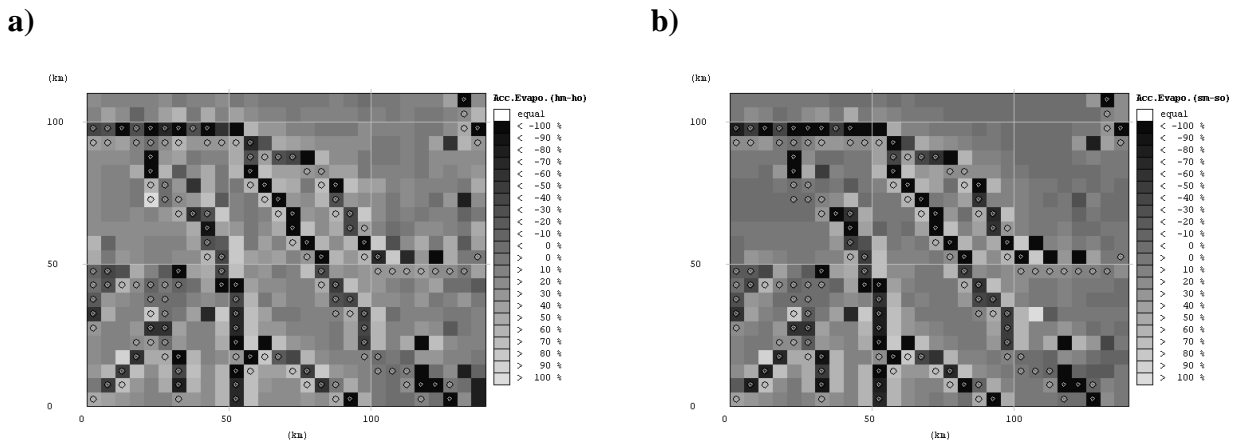


Fig. 4 Like Fig. 2, but for percentage difference of 24 hours accumulated evapotranspiration as obtained by (a) $hm - ho$ scaled by ho , and (b) $sm - so$ scaled by so .

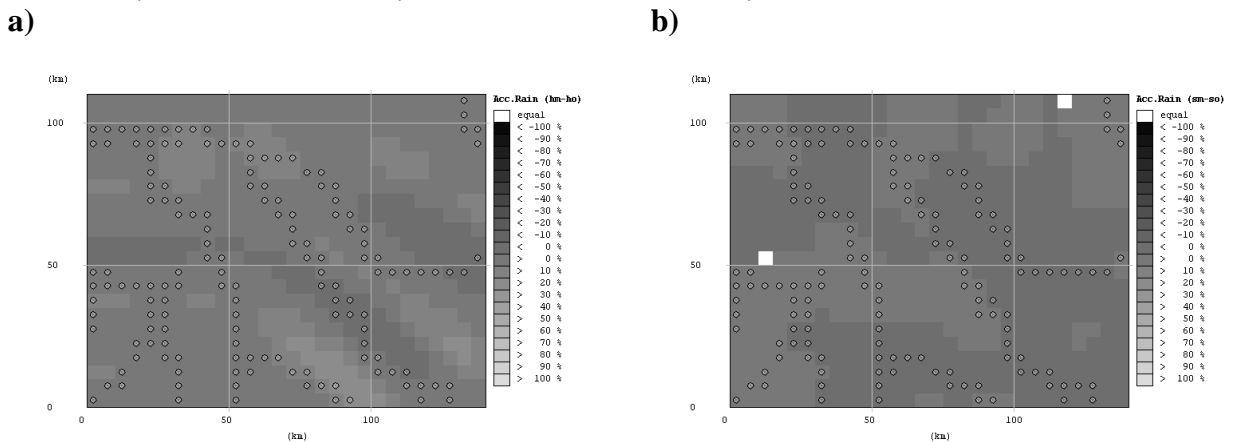


Fig. 5 Like Fig. 2, but for percentage difference of 24 hours accumulated precipitation as obtained by (a) $hm - ho$ scaled by ho , and (b) $sm - so$ scaled by so .

Figure 6 illustrates the soil moisture fluxes in 1 m depth as yielded for the various simulations after 24 hours of integration. Consideration of surface runoff reduces the water fluxes at that

depth, because the water is transported from the location of strong precipitation into the directions of the rivers. Substitution of water meadows to willow-forests leads to increased water fluxes at a depth of 1 m in both cases with and without consideration of surface and channel runoff. In the simulation with consideration of surface runoff, the land-use changes contribute to a stronger change of infiltration, soil water fluxes, and ground water recharge than in that without surface runoff (e.g., Figs. 3, 6). This finding illustrates the need for parameterizations of subgrid-scale surface runoff for investigations on impact of land-use changes on the regional water availability (e.g., Fig. 6).

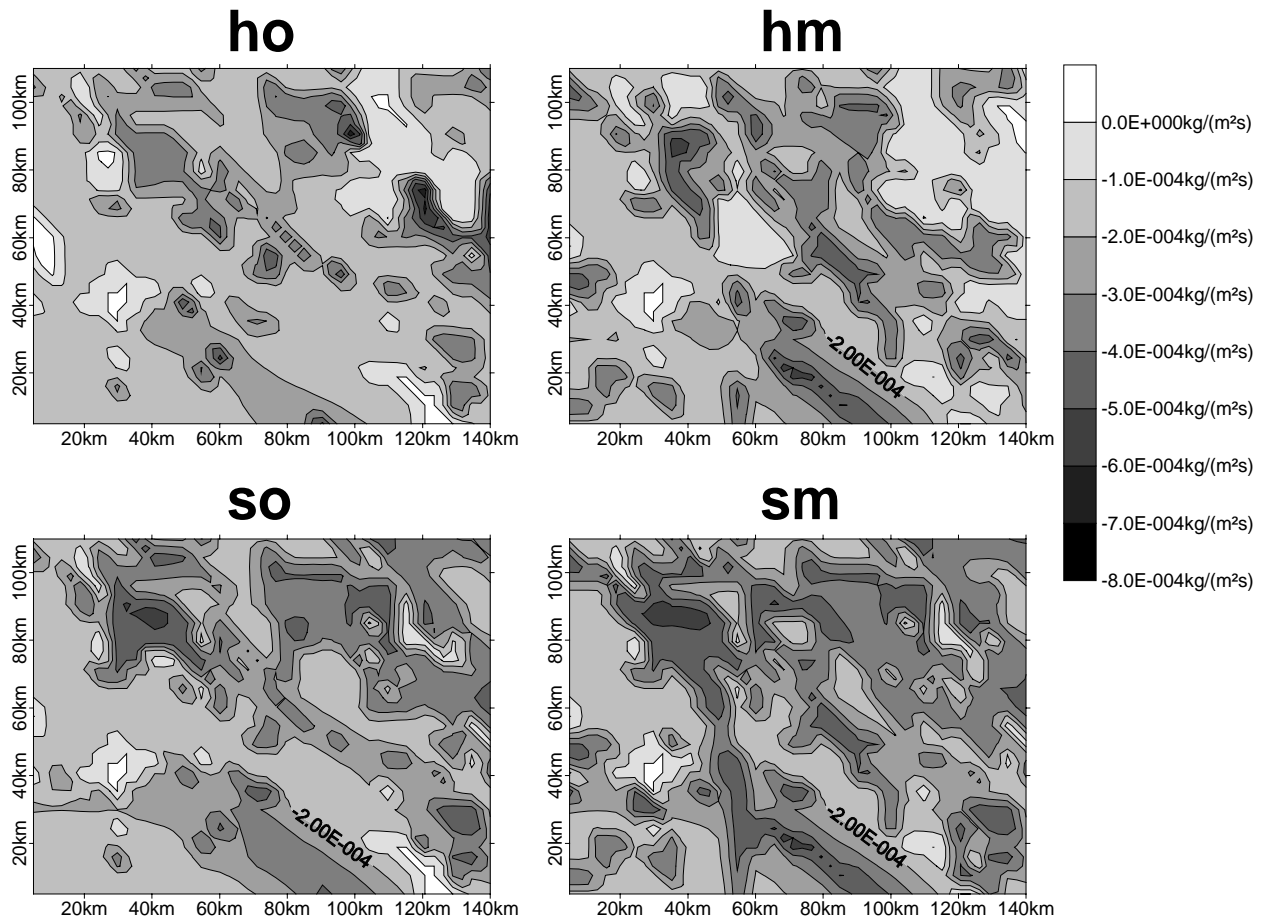


Fig. 6 Soil moisture fluxes at a depth of 1 m after 24 hours of integration as obtained for the simulations using the module for surface and channel runoff without, *ho*, and with land-use changes, *hm*, the simulations without surface and channel runoff module without, *so*, and with land-use changes, *sm*, respectively (from upper left to lower right).

Conclusions and outlook

A module to describe surface and channel runoff within a mesoscale atmospheric model was developed and coupled to a hydro-dynamics soil-vegetation scheme to examine whether these components of the water cycle may play a role for the regional water cycle and prediction of water availability. Simulations with and without these components show that surface runoff as well as land-use changes in favor of willow-forests along the rivers may appreciably influence water fluxes at a depth of 1 m. The sensitivity of water availability to land-use changes was found to be greater when surface runoff is taken into account. Since a huge number of rivers is of subgrid-scale with respect to the grid- or subgrid-resolution of mesoscale atmospheric models, we have to expect that the impact of land-use changes on water availability predicted by mesoscale atmospheric models without consideration of surface runoff may be underesti-

mated if the patch-size of the land-use changes is assumed in a realistic magnitude. Secondary differences in ground water recharge result from the altered surface moisture and evapotranspiration, and, hence, different local recycling of previous precipitation. Based on our findings we may conclude that investigations on the regional water availability and its modification by land-use changes should include surface and channel runoff within the framework of an atmospheric model.

To improve the representation of the water cycle evaporation of water from the river has to be included in the future. Herein, a critical parameter will be the slope of the riverbanks, because the variability of evaporation from the water surface under different water-levels depends on the river surface and, hence, indirectly on the river slope. Unfortunately, such data do not exist for all rivers.

In the future, real cases using measured river inflow at the boundaries have to be performed and evaluated by water-gauges. In doing so, the atmospheric model should be driven by data from a global model to allow long-term studies. Unfortunately, such an evaluation has to be postponed to the future until detailed data sets of river characteristics are also available for small rivers.

Acknowledgements

We would like to express our thanks to K. Fröhlich and K. Friedrich (University Leipzig) for digitizing the land-use data. We thank G. Kramm (MFPA Leipzig), K.E. Erdmann (emc Leipzig), U. Haferkorn (Betriebsanlagen Gesellschaft Brandis), and A. Ziemann (University Leipzig) for fruitful discussion and helpful comments. This study was financially supported by BMBF and DFG under contracts 01LA9849/4 and Mo770/2-1, respectively.

References

- Abbott, M.B., J.C. Bathurst, J.A. Cunge, P.E. O'Connell, J. Rasmussen, 1986. An introduction to the european hydrological system - Système Hydrologique Européen, „SHE“, 2: structure of a physically-based, distributed modelling system. *J. Hydrol.* **87**, 61-77.
- Dingman, S.L., 1993. *Physical Hydrology*. Macmillan Publishing Company, New York, Oxford, Singapore, Sydney, p. 575.
- Eppel, D.P., H. Kapitza, M. Claussen, D. Jacob, W. Koch, L. Levkov, H.-T. Mengelkamp, and N. Werrmann, 1995. The non-hydrostatic mesoscale model GESIMA. Part II: Parameterizations and applications. *Contrib. Atmos. Phys.* **68**, 15-41.
- Graf, W.H., 1998. *Fluvial hydraulics*. John Wiley & Sons, p. 692.
- Hagemann, S., Dümenil, L., 1998. A parameterization of the lateral waterflow for the global scale. *Clim. Dyn.* **14**, 17-31.
- Kapitza, H. and D.P. Eppel, 1992. The non-hydrostatic mesoscale model GESIMA. Part I: Dynamical equations and tests. *Contrib. Phys. Atmos.* **65**, 129-146.
- Kramm, G., 1989. The estimation of surface layer parameters from wind velocity, temperature and humidity profiles by least squares methods. *Bound.-Layer Meteor.* **48**, 315-327.
- Kramm, G., R. Dlugi, N. Mölders, and H. Müller, 1994. Numerical investigations of the dry deposition of reactive trace gases. In: Baldasano, J.M., C.A. Brebbia, H. Power, P. Zannetti (eds.), *Air Pollution II Vol. 1: Computer Simulation*, Computational Mechanics Publications, Southampton, Boston, 285-307.
- Kramm, G., N. Beier, T. Foken, H. Müller, P. Schröder, and W. Seiler, 1996. A SVAT scheme for NO, NO₂, and O₃ - model description. *Meteorol. Atmos. Phys.*, **61**, 89-106.
- Kuhl, S.C., and Miller, J.R., 1992. Seasonal river runoff calculated from a global atmospheric model. *Water Resources Res.* **28**: 2029-2039.
- Lawler, E.A. 1964. Flood routing. In: Chow, V.T. (Ed), *Handbook of applied hydrology; A compendium of water resources technology*, McGraw-Hill Book company, New York.
- Maniak, U., 1993. *Hydrologie und Wasserwirtschaft*. Springer, Berlin, pp. 568.
- Marengo, J.A., Miller, J.R., Russel, G.L., Rosenzweig, C.E. and Abramopoulos, F., 1994. Calculations of river-runoff in the GISS GCM: impact of a new land-surface parameterization and runoff routing model on the hydrology of the Amazon river. *Clim. Dyn.* **10**: 349-361.

- Miller, J.R., Russel, G.L. and Caliri, G., 1994. Continental-scale river flow in climate models. *J. Clim.* 7: 914-928.
- Mölders, N., A. Raabe, and G. Tetzlaff, 1996. A comparison of two strategies on land surface heterogeneity used in a mesoscale β meteorological model. *Tellus* **48A**, 733-749.
- Mölders, N., Raabe, A., 1997. Testing the effect of a two-way-coupling of a meteorological and a hydrologic model on the predicted local weather. *Atmos. Res.* 45: 81-108.
- Mölders, N., G. Kramm, M. Laube, and A. Raabe, 1997. On the influence of bulk-parameterization schemes of cloud microphysics on the predicted water-cycle relevant quantities - a case study. *Met. Zeitschr.* **6**, 21-32.
- Mölders, N., 1999. *Einfache und akkumulierte Landnutzungsänderungen und ihre Auswirkung auf Evapotranspiration, Wolken- und Niederschlagsbildung*. Mitt. Inst. Meteor. Leipzig, **15**, Habilitation thesis (available from the author).
- Mölders, N., Raabe, A., Beckmann, T., 1999a. A technique to downscale meteorological quantities for the use in hydrologic models - Description and first results. *IAHS*, **254**, 89-98.
- Mölders, N., Haferkorn, U., Knappe, S., Döring, J., Kramm, G., 1999b. Evaluation of simulated water budget by means of measurements at Brandis lysimeter station. In: Tetzlaff, G., Grünwald, U. (Eds.), *2. Tagung des Fachausschusses Hydrometeorologie am 15./16. November 1999 in Leipzig*, Wiss. Mitt. Leipzig, **16**, 67-83.
- Mölders, N., U. Haferkorn, S. Knappe, J. Döring, and G. Kramm, 2000. Application of lysimeter- and tensiometer data for evaluation of a module to couple hydrological and atmospheric models. In: Gerold, G. (ed.), *Heterogenität landschaftshaushaltlicher Wasser- und Stoffumsätze in Einzugsgebieten*, Eco Regio, Göttingen, **8**, 97-105.
- Mölders, N., Rühaak, W., 2000. Entwicklung eines Moduls zur Untersuchung von Grundwasserneubildung. Abschlussbericht für den Zeitraum vom 1.11.1998 - 31.12.2000. Förderkennzeichen 01LA9849/4, pp. 85.
- Müller, E., Foken, T., Heise, E. and Majewski, D., 1995. LITFASS - A nucleus for a BALTEX field experiment. *DWD Arbeitsergebnisse* No. 33.
- Orlanski, I., 1976. A simple boundary condition for unbounded hyperbolic flows. *J. Comp. Phys.* **21**, 251-269.
- Pleiss, H., 1977. *Der Kreislauf des Wassers in der Natur*, VEB G. Fischer Verlag, Jena.
- Sausen, R., Schubert, S., Dümenil, L., 1994. A model of the river runoff for use in coupled atmosphere-ocean models. *J. Hydrol.* **155**, 337-352.
- Schröder, W., 1994. *Grundlagen des Wasserbaus: Hydrologie, Hydraulik – Wasserbau*. Werner, Düsseldorf.

Authors' addresses:

Dr. habil. Nicole Mölders, Institut für Meteorologie, Universität Leipzig, Stephanstr. 3, 04103 Leipzig, Germany

Dipl.-Geol. Wolfram Rühaak, Institut für Geowissenschaftliche Gemeinschaftsaufgaben, Sektion Geothermik und Grundwasserhydraulik, Stilleweg 2, 30631 Hannover, Germany

Numerical solution of the Stochastic Collection Equation – Comparison of the Linear Discrete Method and the Method of Moments

Martin Simmel

Abstract

The Linear Discrete Method (LDM; SIMMEL 2000; SIMMEL ET AL. 2000) is used to solve the Stochastic Collection Equation (SCE) numerically. Comparisons are made to the Method of Moments (MOM; TZIVION ET AL. 1999) which is suggested as a reference for numerical solutions of the SCE.

Simulations for both methods are shown for the GOLOVIN kernel (for which an analytical solution is available) and the hydrodynamic kernel after LONG (1974) as it is used by TZIVION ET AL. (1999). Different bin resolutions are investigated and the simulation times are compared. In addition, LDM simulations using the hydrodynamic kernel after BÖHM (1992b) are presented.

The results show that for the GOLOVIN kernel, LDM is slightly closer to the analytic solution than MOM. For the LONG kernel, the low resolution results of LDM and MOM are of similar quality compared to the reference solution. For the BÖHM kernel, only LDM simulations were carried out which show good correspondence between low and high resolution results.

Numerische Lösung der Gleichung für stochastisches Einsammeln – Vergleich der linearen diskreten Methode und der Methode der Momente

Zusammenfassung

Die lineare diskrete Methode (LDM; SIMMEL 2000; SIMMEL ET AL. 2000) wird dazu benutzt, die Gleichung für stochastisches Einsammeln (stochastic collection equation, SCE) numerisch zu lösen. Dabei werden Vergleiche gezogen zur Methode der Momente (Method of Moments, MOM; TZIVION ET AL. 1999), die als Referenz für numerische Lösungen der SCE vorgeschlagen wurde.

Simulationsrechnungen für beide Methoden werden für die Koaleszenzfunktion nach GOLOVIN (für die eine analytische Lösung existiert) und die hydrodynamische Koaleszenzfunktion nach LONG (1974) wie sie von TZIVION ET AL. (1999) verwendet wird, gezeigt. Verschiedene Klassenauflösungen werden untersucht und die Simulationszeiten verglichen. Zusätzlich werden LDM-Simulationen mit der hydrodynamischen Koaleszenzfunktion nach BÖHM (1992b) gezeigt.

Die Ergebnisse für die Koaleszenzfunktion nach GOLOVIN zeigen, daß die LDM der analytischen Lösung etwas näher kommt als MOM. Für die Koaleszenzfunktion nach LONG sind die Ergebnisse von LDM und MOM mit niedriger Auflösung von ähnlicher Qualität verglichen mit der Referenzlösung. Für die Koaleszenzfunktion nach BÖHM wurden nur Simulationen mit der LDM durchgeführt, die eine gute Übereinstimmung der Ergebnisse mit niedriger und hoher Auflösung zeigen.

1 Introduction

Collision and coalescence are important processes in cloud microphysics. They provide the major reason for cloud droplets to grow to rain drops. In bulk models, these processes are parameterized, whereas spectral or bin models are able to solve the governing equation, the stochastic collection equation (SCE), explicitly.

In literature there are various contributions to the numerical solution of the SCE (e. g., KOVETZ AND OLUND 1969; BLECK 1970; BERRY AND REINHARDT 1974; BOTT 1998; TZIVION ET AL. 1999). In addition, mathematical kernels were found, for which the SCE can be solved analytically (GOLOVIN 1963; SCOTT 1968) when certain initial conditions are fulfilled. These analytical solutions are important for the evaluation of the different numerical schemes used in models.

For the hydrodynamic kernel, analytical solutions are not available. Therefore, theoretical considerations and comparisons between different numerical models are a good chance to evaluate the various methods. The high resolution solutions obtained by the Method of Moments (MOM) can be used as reference. This was proposed by TZIVION ET AL. (1999).

2 Equation for the moments

The coalescence process is described by the SCE (e. g. PRUPPACHER AND KLETT 1997). It can be written in the form

$$\begin{aligned} \frac{\partial n(x, t)}{\partial t} = & \frac{1}{2} \int_0^x n(x-y, t)n(y, t)K(x-y, y)dy \\ & - n(x, t) \int_0^\infty n(y, t)K(x, y)dy \end{aligned} \quad (1)$$

where $n(x, t)dx$ is the number of drops in the interval $[x, x+dx]$, measured per unit volume at time t and $K(x, y)$ is the collection kernel. A single drop of mass x belongs to bin k if

$$x_k \leq x < x_{k+1} \quad (2)$$

with $x_{k+1} = px_k$ and $p = const > 1$. In the simulations presented $p = 2, 2^{1/2}, 2^{1/3}, 2^{1/4}$ and $x_1 = \frac{4}{3}\pi r_1^3 \rho_l$ with $r_1 = 1.5625 \mu\text{m}$ and $\rho_l = 1000 \text{ kg m}^{-3}$ are used. The l th moment of the distribution function $n(x, t)$ in bin k is defined as

$$M_k^l = \int_{x_k}^{x_{k+1}} x^l n(x, t) dx \quad (3)$$

After multiplying (1) with x^l and integrating over each bin k , a set of prognostic equations for the moments in each bin k is obtained:

$$\begin{aligned} \frac{dM_k^l(t)}{dt} = & \frac{1}{2} \int_{x_k}^{x_{k+1}} \int_{x_1}^x n(x-y, t)n(y, t)K(x-y, y)dy \\ & - \sum_{j=1}^{J_{MAX}} \int_{x_k}^{x_{k+1}} x^l n(x, t) dx \int_{x_j}^{x_{j+1}} n(y, t)K(x, y)dy \end{aligned} \quad (4)$$

with J_{Max} being the number of bins defined (36, 72, 108, 144 here). The first term on the right-hand side describes the gain of bin k due to collisions that result in drops belonging to bin k (gain term). The second term describes the loss of bin k caused by the collision of a drop of bin k with another drop (loss term). This set of equations is solved

case	golo1	golo3	long1	long3	böhm1	böhm3
M_0 (g/m ³)	1	3	1	3	1	3
x_0 (kg)	$3.33 \cdot 10^{-12}$	10^{-11}	$3.33 \cdot 10^{-12}$	10^{-11}	$3.33 \cdot 10^{-12}$	10^{-11}
kernel	GOLOVIN	GOLOVIN	LONG	LONG	BÖHM	BÖHM
time (min.)	20, 40	10, 15	20, 40	10, 15	20, 40	10, 15

Table 1: Overview of all cases chosen. Two different initial distributions are combined with three kernels. For each case, several runs with different methods (MOM, LDM) and varying resolution are performed.

numerically by the Linear Discrete Method (LDM; SIMMEL 2000; SIMMEL ET AL. 2000). The results are compared to the established and accurate MOM (TZIVION ET AL. 1987, 1999). For further information about the numerical methods used the reader is referred to the literature cited above.

3 Results and discussion

We present simulations using the numerical methods mentioned in the previous section. We conducted simulations for different resolutions ($p = 2$, $p = 2^{1/2}$, $p = 2^{1/3}$, and $p = 2^{1/4}$) for the GOLOVIN kernel (GOLOVIN 1963) and for the hydrodynamic kernels after LONG (1974) and BÖHM (1992b), respectively. As initial distribution we use the same exponential function as TZIVION ET AL. (1987, 1999)

$$n(x) = 4 \frac{N_0 x}{x_0^2} \exp[-2x/x_0] \quad , \quad (5)$$

with $N_0 = 3 \cdot 10^8 \text{ m}^{-3}$ and either $x_0 = 3.33 \cdot 10^{-12} \text{ kg}$ or $x_0 = 10^{-11} \text{ kg}$ (see Table 1). This corresponds to liquid water contents of 1 g m^{-3} and 3 g m^{-3} , respectively. Table 1 gives an overview over all runs performed and the parameters used. For all runs the water mass is conserved (except numerical inaccuracies).

3.1 Golovin kernel

Using the GOLOVIN kernel, the SCE can be solved analytically for various initial distributions (GOLOVIN 1963, SCOTT 1968). The GOLOVIN or "sum of mass" kernel is given as

$$K(x, y) = b(x + y) \quad (6)$$

with $b = 1.5 \text{ m}^3 \text{ s}^{-1} \text{ kg}^{-1}$ and x, y the mass of the colliding drops. Out of all kernels for which an analytical solution of the SCE is known, the GOLOVIN kernel is closest to the hydrodynamical kernel and therefore is a good test for the numerical method used.

Results Figure 1 shows the good correspondence of the numerical LDM simulations for all resolutions with the analytical solution. `golo1` is on the left, `golo3` on the right. As expected, numerical solutions are closer to the analytical solution for higher resolution in both cases, but even for 36 bins, the LDM solution is pretty good.

Figure 2 compares the mass distribution function for LDM and MOM with 72 bins each to the analytical solution. LDM is closer to the analytical solution than MOM for both cases (`golo1` on the left and `golo3` on the right), but generally, they are in good agreement with the analytical solution. The drop growth tends to be overestimated by both methods for the GOLOVIN kernel.

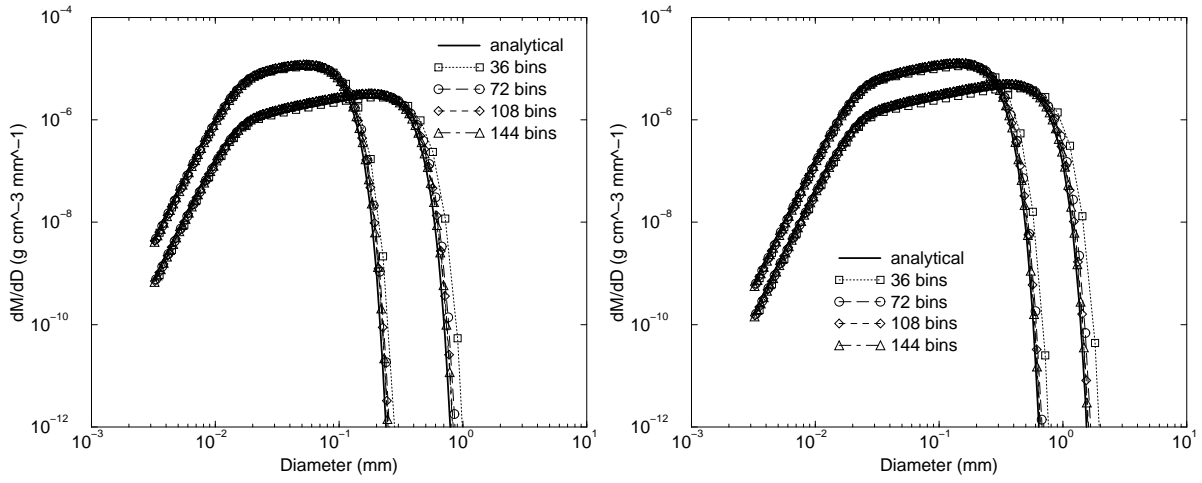


Figure 1: *Mass distribution for the GOLOVIN kernel using LDM with different resolutions compared to the analytical solution. Results are shown for the golo1 (left) and golo3 (right) cases, time step was 1 s.*

3.2 Hydrodynamic kernel

The hydrodynamic kernel of two interacting drops with masses x and y and radii r_x and r_y , respectively, is defined as

$$K(x, y) = \pi(r_x + r_y)^2 E(x, y) |v_x - v_y| \quad (7)$$

with the terminal fall velocities v_x and v_y and the collection efficiency $E(x, y)$, given as

$$E(x, y) = E_{coll}(x, y) E_{coal}(x, y) \quad . \quad (8)$$

Here, $E_{coll}(x, y)$ is the collision efficiency and $E_{coal}(x, y)$ the coalescence efficiency. In literature, different data from observations and theoretical investigations can be found. We choose one kernel after LONG (1974) and one after BÖHM (1992b). Using the first one, we can compare our results with those of TZIVION ET AL. (1999). The second one based more on theoretical considerations using boundary layer theory. It is valid

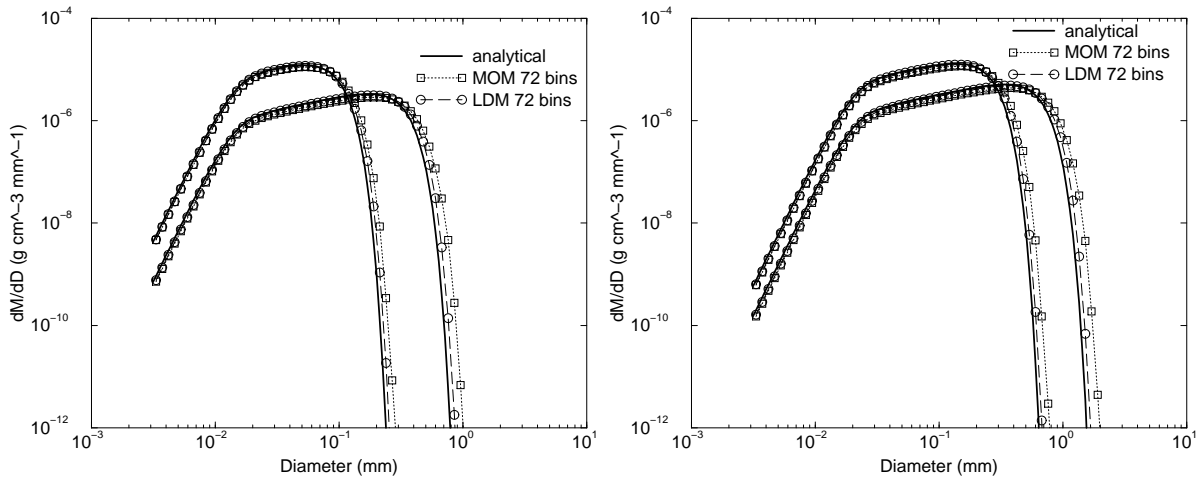


Figure 2: *Mass distribution for the GOLOVIN kernel using LDM and MOM with $p = 2^{1/2}$ compared to the analytical solution. Results are shown for the golo1 (left) and golo3 (right) cases, time step was 1 s.*

Diameter d_x (in μm)	α	β
< 79.37	$4.5795 \cdot 10^5$	$2/3$
$79.37\text{--}800$	$4.962 \cdot 10^3$	$1/3$
$800\text{--}4031.74$	$1.732 \cdot 10^3$	$1/6$
> 4031.74	$9.17 \cdot 10^2$	0

Table 2: Coefficients for the velocity calculation following eq. (10).

for hydrometeors that are regarded as porous spheroids moving in the orientation which offers the maximum drag to motion (BÖHM 1992a). The particles are described by four parameters: particle mass, semi axis of symmetry, equatorial radius, and the ratio of the effective over the circumscribed cross-sectional area. Therefore, the BÖHM theory can be applied not only to water drops but as well to different shapes of ice particles. This would lead to different kernels for the interaction of ice particles with drops or other ice particles.

3.2.1 Long kernel

The collision efficiency is approximated by

$$E_{col}(x, y) = \begin{cases} 0 & \text{for } r_x \leq 3\mu\text{m} \\ 1 & \text{for } r_y \geq 50\mu\text{m} \\ 4.5 \cdot 10^4 r_y^2 (1 - 3 \cdot 10^{-4}/r_x) & \text{else} \end{cases} \quad (9)$$

The coalescence efficiency is assumed to be unity ($E_{coal}(x, y) = 1$). The collection kernel is calculated via eqs. (7) and (8) with the terminal velocity

$$v_x = \alpha x^\beta \quad (10)$$

using the coefficients α and β of Table 2. For all calculations concerning the LONG kernel cgs-units have to be used.

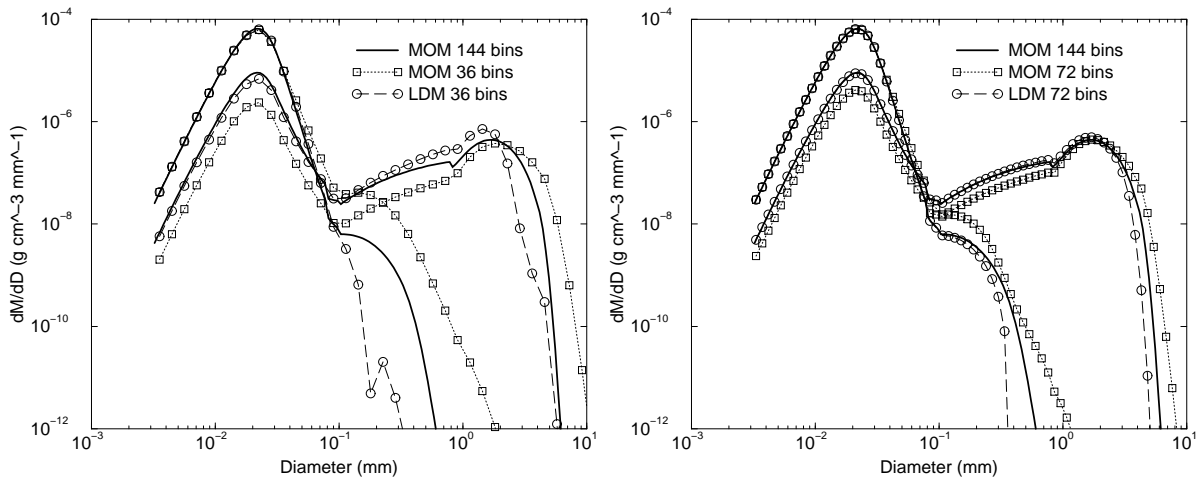


Figure 3: Mass distribution for the LONG kernel using LDM and MOM with $p = 2$ (left) and $p = 2^{1/2}$ (right) compared to the reference solution (MOM with $p = 2^{1/4}$). Results are shown for the long1 case, time step was 1 s.

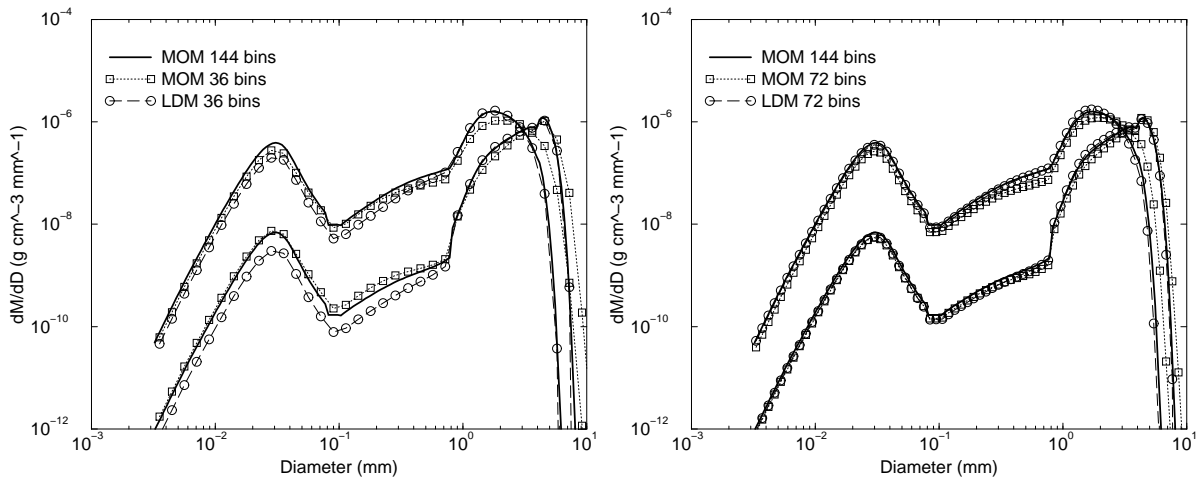


Figure 4: Mass distribution for the LONG kernel using LDM and MOM with $p = 2$ (left) and $p = 2^{1/2}$ (right) compared to the reference solution (MOM with $p = 2^{1/4}$). Results are shown for the long3 case, time step was 1 s.

For the hydrodynamic kernel an analytic solution does not exist. Therefore, we use the MOM-solution with 144 bins ($p = 2^{1/4}$) as reference solution as it is proposed by TZIVION ET AL. (1999).

Results Figure 3 shows the results for long1 using MOM and LDM with $p = 2$ (left) and $p = 2^{1/2}$ (right) compared to the reference solution. For both resolutions, MOM tends to overestimate the growth, whereas LDM seems to underestimate the mass of the larger drops.

Figure 4 is the same but for long3. The deviations for long3 are much smaller than for long1. Possibly, both numerical methods are not so sensitive to variations of the kernel for larger drops. The shorter overall integration time of long3 is not the reason. The results are more accurate when the distributions are shifted to larger drops (compare 20 min. to 40 min. for long1 and 10 min. to 15 min. for long3).

Compared to the GOLOVIN kernel, the errors for the hydrodynamic kernel are larger for both initial distributions and all resolutions. Nevertheless, we can state that if a model shows good results for the GOLOVIN kernel it will give good results for the hydrodynamic kernel as well.

3.2.2 Böhm kernel

The hydrodynamic kernel after BÖHM (1992b) is a semiempirical solution based on boundary-layer theory. It is in good agreement with measured coalescence efficiencies and other theories (Stokes', modified Oseen, superposition method). For drops of equal size and terminal velocity, non-vanishing kernels are predicted.

Results Figure 5 shows the results for the cases böhm1 and böhm3 for different resolutions using LDM. Only a slight retardation can be seen for low resolutions compared to the 144 bin-solution. Especially for böhm3, the results for the different resolutions are very close to each other. This supports the assumption that the growth process can be predicted more accurate for larger drops.

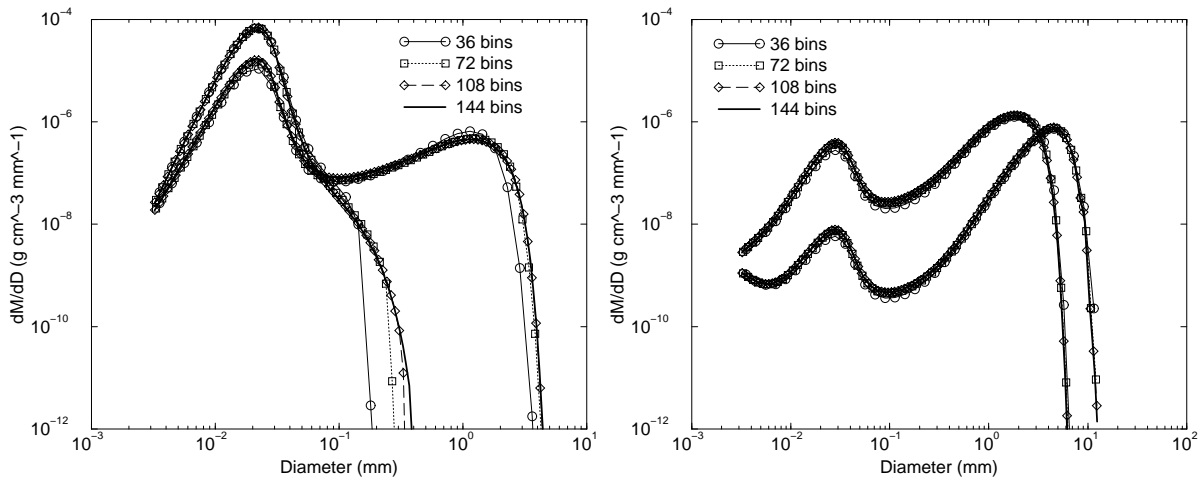


Figure 5: *Mass distribution for the BÖHM kernel using LDM with different resolutions for the cases böhm1 (left) and böhm3 (right) with the averaging procedure. Time step was 1 s.*

3.3 Computation times

Table 3 shows the computation times for the runs with the GOLOVIN (`golo1` and `golo3`) and the LONG (`long1` and `long3`) kernel. Figure 6 shows the same for the `long1` and `golo1` cases on the left and the `golo3` and `long3` cases on the right. For high resolutions ($p = 2^{1/3}$, $p = 2^{1/4}$), LDM is much faster than MOM, whereas for low resolutions ($p = 2$), LDM seems to be somewhat slower, depending on the case. For MOM, `golo1` and `long1` need almost the same computation time, as well as `golo3` and `long3`. For LDM, the GOLOVIN cases are faster than the corresponding LONG cases (`golo1` vs. `long1` and `golo3` vs. `long3`).

Generally, empty bins are not taken into account to save computation time. Therefore, underestimation of the growth saves computation time whereas overestimation leads to longer computation times. These effects should be negligible.

4 Conclusions

A new spectral model LDM for a numerical solution of the SCE was compared to the well-known MOM whose 144-bin numerical solution was proposed as a reference for the numerical solution of the SCE.

For different kernels, simulations with varying resolution were carried out. For the GOLOVIN kernel, the results were compared to the analytical solution, too. For the hydrodynamic kernel, no analytical solution is available. For the BÖHM kernel, the solutions

	$p = 2$		$p = 2^{1/2}$		$p = 2^{1/3}$		$p = 2^{1/4}$	
	MOM	LDM	MOM	LDM	MOM	LDM	MOM	LDM
<code>golo1</code>	2.62 s	2.46 s	22.28 s	9.57 s	63.28 s	20.18 s	159.18 s	36.81 s
<code>golo3</code>	1.04 s	1.59 s	7.79 s	4.60 s	22.76 s	11.82 s	62.10 s	15.63 s
<code>long1</code>	2.54 s	4.61 s	22.32 s	17.98 s	62.02 s	37.44 s	169.53 s	57.88 s
<code>long3</code>	1.00 s	2.68 s	7.85 s	7.45 s	21.75 s	16.33 s	60.19 s	27.06 s

Table 3: *Run times of the runs performed on an IBM RS6000. For `golo1` and `long1` the integration time was 40 min., for `golo3` and `long3` it was 15 min.*

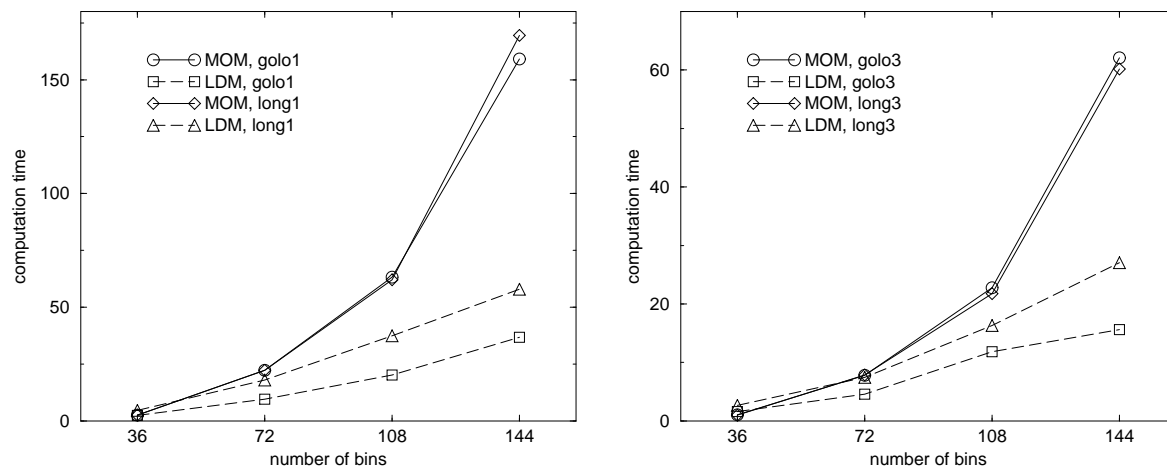


Figure 6: Run times for the cases *golo1* and *long1* on the left and *golo3* and *long3* on the right for MOM (solid line) and LDM (dashed line).

using different resolution are very close to each other which means that LDM gives good results even when using low resolution.

Especially for high resolutions, LDM needs less computation time than MOM without being less accurate. Considering both, accuracy and integration time, LDM is a good choice for the numerical solution of the SCE and therefore, for a spectral model.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under contracts TE 51/11-1 and TE 51/11-3. Thanks to Prof. Z. Levin and Dr. T. Reisin (Tel Aviv University, Israel) for the numerical code of the MOM which was used for the comparisons.

References

- BERRY, E. X. AND R. L. REINHARDT, 1974. An analysis of cloud drop growth by collection. Part I: Double distributions. *Journal of the Atmospheric Sciences*, **31**, 1814–1824.
- BLECK, R., 1970. A fast, approximative method for integrating the stochastic coalescence equation. *Journal of Geophysical Research*, **75**, 5165–5171.
- BÖHM, J. P., 1992a. A general hydrodynamic theory for mixed-phase microphysics: Part I: Drag and fall speed of hydrometeors. *Atmospheric Research*, **27**, 253–274.
- BÖHM, J. P., 1992b. A general hydrodynamic theory for mixed-phase microphysics: Part II: Collision kernels for coalescence. *Atmospheric Research*, **27**, 275–290.
- BOTT, A., 1998. A flux method for the numerical solution of the stochastic collection equation. *Journal of the Atmospheric Sciences*, **55**, 2284–2293.
- GOLOVIN, A. M., 1963. The solution of the coagulation equation for cloud droplets in a rising air current. *Bull. Acad. Sci. USSR Geophys. Ser.*, **5**, 783–791.
- KOVETZ, A. AND B. OLUND, 1969. The effect of coalescence and condensation on rain formation in a cloud of finite vertical extent. *Journal of the Atmospheric Sciences*, **26**, 1060–1065.

- LONG, A. B., 1974. Solutions to the droplet coalescence equation for polynomial kernels. *Journal of the Atmospheric Sciences*, **31**, 1040–1052.
- PRUPPACHER, H. R. AND J. D. KLETT, 1997. *Microphysics of clouds and precipitation*. Kluwer Academic Publisher.
- SCOTT, W. T., 1968. Analytic studies of cloud droplet coalescence. *Journal of the Atmospheric Sciences*, **25**, 54–65.
- SIMMEL, M., 2000. Two numerical solutions for the stochastic collection equation. *Wissenschaftliche Mitteilungen aus dem Institut für Meteorologie der Universität Leipzig und dem Institut für Troposphärenforschung e. V. Leipzig*, **17**, 61–73.
- SIMMEL, M., N. MÖLDERS, AND G. TETZLAFF, 2000. A numerical method for the solution of the sce using two prognostic moments. Volume 1, pp. 489–492. 13. International Conference on Clouds and Precipitation.
- TZIVION, S., G. FEINGOLD, AND Z. LEVIN, 1987. An efficient numerical solution to the stochastic collection equation. *Journal of the Atmospheric Sciences*, **41**, 1648–1661.
- TZIVION, S., T. G. REISIN, AND Z. LEVIN, 1999. A numerical solution of the kinetic collection equation using high spectral grid resolution: A proposed reference. *Journal of Computational Physics*, **148**, 527–544.

Address of the author:

Martin Simmel
Institut für Meteorologie
Universität Leipzig
Stephanstr. 3
04103 Leipzig
smartie@meteo.uni-leipzig.de

Berücksichtigung atmosphärischer Schallausbreitungsbedingungen beim Lärmschutz

A. Ziemann, K. Arnold und A. Raabe

Zusammenfassung:

Neben der Abnahme des Schalldruckes bzw. der Schallintensität mit zunehmender Entfernung von einer Schallquelle (Kugelwellendivergenz) spielt bei der Schallausbreitung in der atmosphärischen Grenzschicht der Einfluss der höhenveränderlichen meteorologischen Größen, speziell der Temperatur und des Windvektors, auf die Geschwindigkeit und Richtung der Schallwellen eine entscheidende Rolle.

Da in den bisherigen Richtlinien zur Erstellung von Lärmkatastern (über einen bestimmten Zeitraum gemittelte Schalldruckpegelverteilung in einem Gebiet) oder zur Messung der Schallimmission in der Umgebung von Schallquellen aktuelle meteorologische Informationen nicht einfließen, wird in dieser Studie untersucht, inwieweit ein entsprechend erweitertes Konzept bzw. Modell zu modifizierten Ergebnissen führt.

Die Untersuchungen münden in ein Verfahren, das die Schallausbreitungsbedingungen an einem Ort für einen einzelnen Zeitpunkt bzw. während eines bestimmten Zeitraumes entsprechend der aktuellen thermischen Atmosphärensichtung und den Vertikalprofilen von Windgeschwindigkeit und –richtung bestimmt. Die dafür benötigten Eingangsinformationen können durch Atmosphärenmodelle jederzeit und für beliebige Gitterpunkte bereitgestellt werden.

Summary:

Except for the decrease of the acoustic pressure and the acoustic intensity with increasing distance between sound source and receiver (spherical divergence) the sound propagation and thereby the velocity and direction of sound waves inside the atmospheric boundary layer will be mainly influenced by the height-variable meteorological quantities, especially temperature and wind vector.

Because actual meteorological information are ignored by the existing rules to create a noise register (averaged sound-pressure level distribution during a definite time interval and over a fixed region) or for the measurement of noise in the environment of sound sources, it will be investigated in this study whether an extended concept and model for the sound propagation will lead to modified results or not.

The investigations lead to a technique which quantify the local conditions of sound propagation for one time or for a definite period corresponding to the vertical profiles of the air temperature as well as of wind speed and wind direction. The needed input data can be provided anytime and for any grid point by an atmospheric model.

1. Einleitung

Eine wichtige Aufgabe des Umweltschutzes besteht nach dem Bundesimmissionsschutzgesetz (BImSchG, 1992) in der Überwachung von Geräuschimmissionen.

Die Grenzen der dabei bisher verwendeten Modelle (s. Hirsch und Buchta, 1993), die auf der TA Lärm (1968, 1998) basieren, zeigen sich nach einem Vergleich mit Messwerten vor allem darin, dass die verwendeten Verfahren die atmosphärischen Einflüsse (Vertikalprofile von Temperatur und Wind) auf die Schallausbreitung unberücksichtigt lassen (s. Kühner, 1993). Auch Modellsimulationen, welche eine Schallstrahlenrefraktion bei der Berechnung der

Schalldruckpegelverteilungen berücksichtigen, weisen deutlich darauf hin, dass Temperatur- und Windprofil einen großen Einfluss auf die Schallausbreitung ausüben (s. Modell LARKHILL, Turton et al., 1988a, b).

Da die vorhandenen Modelle jedoch entweder die meteorologischen Umgebungsbedingungen unzureichend beinhalten oder im operationellen Dienst bei der Herstellung von Lärmkatastern bzw. der aktuellen Prognose von Schallausbreitungsbedingungen nicht einsetzbar sind, wird in dieser Studie (s. Raabe et al., 2000) ein neues Konzept bzw. Modell zur Einbeziehung des Atmosphärenzustandes in die Lärmprognose vorgestellt.

2. Einfluss meteorologischer Größen auf die Schallausbreitung

Abgesehen von der Abnahme des Schalldruckes bzw. der Schallintensität mit zunehmender Entfernung von der Schallquelle (sphärische Divergenz) spielen bei der Schallausbreitung der Einfluss von Inhomogenitäten in der Atmosphäre auf die Geschwindigkeit und Richtung der Schallwellen einschließlich Refraktion durch Wind- und Temperaturgradienten sowie die Schallstreuung an Turbulenzelementen, die Luftabsorption und die Schalldämpfung durch den Erdboden eine Rolle (s. z.B. Piercy et al., 1977; Embleton, 1996). Diese Erscheinungen werden vor allem durch die Frequenz der ausgesendeten Schallwelle, die Sender-Empfänger-Geometrie, die meteorologischen Größen und ihre Fluktuationen (z.B. auch Vertikalgradienten) sowie die Eigenschaften der Oberfläche beeinflusst.

Im folgenden wird eine Zusammenfassung der wichtigsten Ergebnisse zum Einfluss der räumlich und zeitlich variablen meteorologischen Größen Temperatur und Windvektor gegeben, da infolge von Gradienten dieser Größen eine Schallstrahlenrefraktion in der Atmosphäre verursacht wird, die mit dem hier vorgestellten LIM-Schallstrahlenmodell untersucht werden soll.

Vorhandene Wind- und Temperaturgradienten sind die Ursache für die Refraktion von Schallstrahlen in der Atmosphäre und damit die Änderung der Schallausbreitungsrichtung. Bereits ab einer Entfernung von wenigen Dekametern von der Schallquelle hat die Refraktion von Schallstrahlen merkliche Auswirkungen auf die Schallausbreitung und eine damit verbundene Änderung der Schalldruckamplitude (s. Hallberg et al., 1985). Diese Wirkung erhöht sich mit zunehmender Entfernung von der Schallquelle. Infolge der Schallstrahlenrefraktion kann es so zu einer relativen Verstärkung oder Abschwächung im Vergleich zu einer homogenen Atmosphäre ohne vertikale Unterschiede in den meteorologischen Größen kommen.

Der gekoppelte Einfluss von vertikalen Wind- und Temperaturprofilen auf die Schallausbreitung in der Atmosphäre wurde u.a. bereits von Wiener und Keast (1959) sowie Pridmore-Brown (1962) untersucht. Dabei zeigte sich deutlich, dass zu dem isotrop wirkenden Einfluss einer veränderlichen Lufttemperatur der richtungsabhängige Einfluss des Windvektors hinzukommt. Die Schallstrahlen werden im Mitwindfall (bei vorausgesetzter Zunahme der Windgeschwindigkeit mit der Höhe) in Richtung Erdoberfläche und bei Gegenwind aufwärts gebrochen. Bei einer Schallausbreitung entgegen der Windrichtung entsteht entsprechend den theoretischen Berechnungen mit Hilfe der geometrischen Akustik eine Schallschattenzone, deren Auswirkung auf den Schalldruckpegel in der Realität jedoch durch Oberflächenwellen entlang der Erdoberfläche gemildert wird (s. Ostashev, 1997). Eine Abnahme der Temperatur mit der Höhe verschärft die Aufwärtsrefraktion bei Gegenwind und schwächt die abwärts gerichtete Refraktion im Mitwindfall ab. Eine Temperaturinversion übt einen entgegengesetzten Effekt aus.

Das Auftreten von Wind- und Temperaturgradienten in der Atmosphäre führt dementsprechend nicht nur zu einer Veränderung der Ausbreitungsrichtung von Schallstrahlen, sondern in Verbindung damit auch zu einer Schalldämpfung (s. Tabelle 1).

Tab. 1: Auswirkungen der Schallstrahlenrefraktion durch vertikale Temperatur- und Windgradienten (gekoppelter Einfluss: Änderung der Schallstrahlenrichtung und Bodeneinfluss) – eine Auswahl. S...Sender, R...Empfänger, S-R...Sender-Empfänger-Entfernung.

Quelle	S- und R-Höhe	Maximale zusätzliche Schalldruckpegeldämpfung
Embleton, 1996 (Messungen)	S: in Bodennähe R: 2 m	- Inversionen bei S-R: 4 km: 9-12 dB - Temperaturabnahme+Gegenwind bei S-R: 230 m: 25-40 dB bei 250 Hz-1 kHz
Foss, 1979 (Messung)	S: 1,2 m R: 0-4,3 m	- Unterschied: 13 dB zwischen Mit- und Gegenwindausbreitung bei 1-3 kHz
Kriebel, 1971 (analytische Berechnung)	S und R am Boden	- 11 dB bis zum Erreichen der Schallschattenzone
L'Espérance et al., 1993 (Messung und Modell)	S : 1,8 m R : 1,75 m	- Mitwind+Temperaturinversion (nachts): 2 dB - Gegenwind+Temperaturabnahme (tags): 32 dB bei einer S-R-Entfernung von 350 m
Lamancusa und Daroux, 1996 (Modellierung)	Keine Angaben	- 10 dB-Dämpfung bei $1,2 \text{ ms}^{-1}$ Windgeschwindigkeit (nur Refraktion)
Wiener und Keast, 1959 (Messung)	S: 3,7 m R: 1,5 m	- Gegenwind (+labil): 25-30 dB bei 500 m-1400 m S-R - Seitenwind (+neutral): 15 dB bei 250 m-1400 m S-R - Labil+Mitwind: keine Schattenzone (Temperatureffekt > Windeffekt) - Stabil+schwacher Wind (nachts): keine Schattenzone

3. Entwicklung des LIM-Schallstrahlenmodells

3.1 Überblick zur Schallstrahlenmodellierung

Die Ausbreitung des Schalls kann mit Schallstrahlen beschrieben werden, wenn die Modellvorstellungen der geometrischen Akustik anwendbar sind (kleine Wellenlängen im Vergleich zu den Ausmaßen des Untersuchungsobjektes bzw. Wellenlänge groß gegenüber dem Gradienten des Refraktionsindex). Schallstrahlen stellen dann Linien dar, deren Tangenten mit der Ausbreitungsrichtung der Schallwelle übereinstimmen. Obwohl die Anwendung der Schallstrahlentheorie mit Einschränkungen verbunden ist, bietet sie gegenüber aufwendigen Wellenmodellen folgende Vorteile: leichte Visualisierung der Schallausbreitung in der Atmo-

sphäre und einfache Berücksichtigung von Inhomogenitäten im Medium durch Anwendung des Brechungsgesetzes.

Bei Verwendung linearer vertikaler Profile für die Schallgeschwindigkeit im unbewegten Medium und unter Annahme horizontaler Homogenität können Schallstrahlen als Kreisbögen (s. auch VDI 2714, 1988) beschrieben werden, deren Radien abhängig von der vertikalen Temperaturverteilung sind (Sager, 1974; Franck und Sager, 1974). Bei Abnahme der Temperatur mit der Höhe sind die Schallstrahlen vom Boden an aufwärts gekrümmt, der horizontal von der Schallquelle abgehende Schallstrahl begrenzt dabei eine unter dem Strahl liegende Schattenzone. Bei Temperaturinversionen werden die Schallstrahlen im Gegensatz dazu zur Erdoberfläche hin gebrochen.

Einen Überblick zu Schallausbreitungsmodellen, die auf der Wellentheorie und der geometrischen Akustik beruhen und die meteorologischen Ausbreitungsbedingungen beachten, geben u.a. Birkhoff (1988) und Ostashev (1997).

3.2 Grundlagen des LIM-Modells

Die vorgestellten Modelle in der Literatur sind jedoch entweder rechenintensiv und nicht operationell anwendbar oder beinhalten die meteorologischen Einflüsse auf die Schallausbreitung nur in vereinfachter Weise (Basis: spezielle Schallgeschwindigkeitsprofile aus einfachen Temperatur- oder Windprofilen). Das im folgenden vorgestellte Schallstrahlenmodell ist demgegenüber effektiv und operationell anwendbar, sofern die benötigten (meteorologischen) Eingangsdaten z.B. aus Wettervorhersagemodellen zur Verfügung stehen.

Ausgehend vom bekannten Snelliusschen Refraktionsgesetz für unbewegte Medien (s. Pierce, 1989) kann ein verallgemeinertes Brechungsgesetz für geschichtete Medien (Temperatur und Windgeschwindigkeit nur von der Höhe abhängig) abgeleitet werden (s. Li et al., 1998; Ostashev, 1997).

Bei schräg zur Windrichtung verlaufender Schallausbreitung müssen die Geschwindigkeiten der Medien v_1 und v_2 sowie die Machzahlen Ma_1 und Ma_2 als Vektorgrößen eingeführt werden. Im Refraktionsgesetz muss in diesem Fall die in Schallausbreitungsrichtung liegende Komponente $Ma_1 \cos \varphi$ bzw. $Ma_2 \cos \varphi$ (φ ist der Winkel zwischen Wind- und Schallausbreitungsrichtung) verwendet werden (siehe Wöhle, 1984; Ziemann, 2000).

$$\sin \alpha_2 = \frac{k_1 \sin \alpha_1}{k_2 + (k_2 |\vec{Ma}_1| \cos \varphi - k_1 |\vec{Ma}_2| \cos \varphi) \sin \alpha_1}$$

mit $|\vec{Ma}_1| = \frac{|\vec{v}_1|}{c_2}$, $|\vec{Ma}_2| = \frac{|\vec{v}_2|}{c_1}$ (3.1)

Die Variablen k_1 und k_2 bezeichnen die Wellenzahlen der Medien an einer Schichtgrenze. Besonders bei höheren Windgeschwindigkeiten muss entsprechend Gleichung (3.1) der Windvektor Beachtung finden, um Fehler in der Schallausbreitungsberechnung zu vermeiden. Ein Nachteil der vereinfachenden zweidimensionalen Betrachtungsweise ist jedoch, dass bei einer Windrichtung senkrecht zur Schallausbreitung der Wind keinen Einfluss hat, obwohl sich der tatsächlich durchlaufene Schallweg durch eine Krümmung der Schallstrahlen verlängert. Zur Lösung dieses Problems müsste eine dreidimensionale Betrachtung bzw. Schallstrahlen-Modellierung angewendet werden. In der praktischen Anwendung ist die Auswirkung einer Windrichtung quer zur Schallausbreitung als relativ gering einzuschätzen, da durch den räumlichen Versatz der Schallstrahlen diese durch veränderte Temperatur- und Windprofile (bedingt durch die Inhomogenität meteorologischer Felder wegen Unterschieden in den Un-

tergrundbedingungen) zum Boden bzw. zu höheren Atmosphärenschichten hin gebrochen werden, ohne den Schallpegel am Immissionsort zu beeinflussen.

Den zwischen 2 Höhenniveaus zurückgelegten Schallweg $\Delta s = \sqrt{(\Delta x)^2 + (\Delta z)^2}$ in einer horizontal homogenen Atmosphäre mit beliebigen Temperatur- und Windvektorprofilen erhält man unter Verwendung von (3.1) und

$$\tan \alpha_2 = \frac{\sin \alpha_2}{\sqrt{1 - \sin^2 \alpha_2}} = \frac{\Delta x}{\Delta z} \quad (3.2)$$

für geringmächtige Schichten der Atmosphäre, in denen Ma und k als konstant betrachtet werden können. Mit dieser Approximation können nicht nur einfache analytische, sondern auch komplexere Profile meteorologischer Größen als Eingangsdaten für die Schallausbreitungsrechnung verwendet werden.

Die Atmosphäre wird dabei in Höhenniveaus der variabel vorgebbaren Dicke Δz eingeteilt, die in horizontaler Richtung parallel zur Erdoberfläche verlaufen. Je kleiner die Abschnitte Δz gewählt werden, desto genauer kann der Schallstrahlenverlauf bestimmt werden.

Grundsätzlich breiten sich die Schallstrahlen mit einem Emissionswinkel nahe 90° (parallel zur Erdoberfläche) wesentlich horizontal geradliniger aus als Strahlen mit geringeren Abgangswinkeln. Die Schallstrahlenausbreitung mit großen Emissionswinkeln ist außerdem sehr sensitiv auf die meteorologischen Bedingungen (siehe Abb. 1 und Abb. 2), d.h., bei einer Schallstrahlenmodellierung ist der Winkelbereich um 90° besonders stark aufzulösen, da bereits kleinste Änderungen der Werte zu großen Änderungen in der Reichweite eines Schallstrahles führen können.

Um die Vergleichbarkeit der Ergebnisse mit anderen Schallstrahlenmodellen zu demonstrieren, sind in den Abbildungen 1 und 2 Beispiele für eine Schallausbreitungsmodellierung unter Einfluss eines vertikalen Temperaturgradienten ohne Betrachtung eines Windeinflusses dargestellt.

Bei dieser Simulation wurden zwei einfache lineare Temperaturprofile als Eingangsdaten für das Schallstrahlenverfolgungsmodell verwendet.

Als Abstrahlwinkel von der Schallquelle (in 1 m Höhe) wurde ein Bereich von 80° bis 89° (Winkel gegen das Einfallslot des Schallstrahls) gewählt. Ein Schallstrahl kann entlang eines Wertes für den Emissionswinkel in der horizontalen (Entfernung vom Sender) bzw. vertikalen (Höhe) Ausbreitung verfolgt werden.

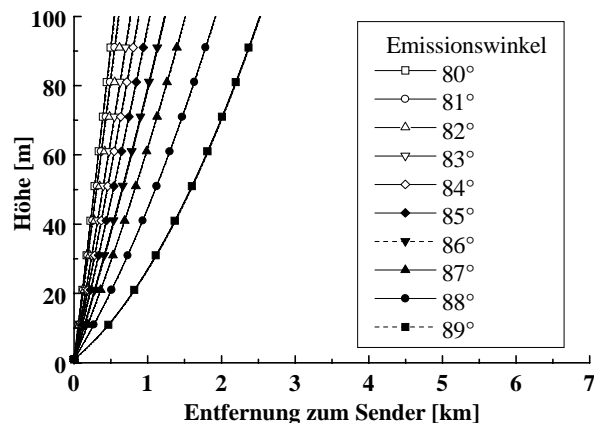


Abb. 1: Schallstrahlenverlauf (Emissionswinkel: $80^\circ \dots 89^\circ$) für eine Lufttemperaturabnahme um $1 \text{ K}/100 \text{ m}$.

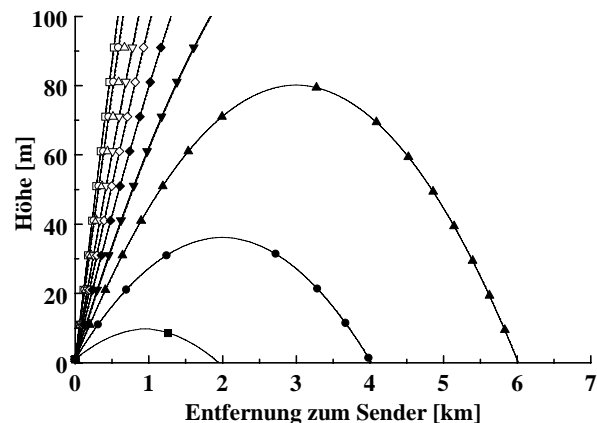


Abb. 2: Schallstrahlenverlauf (Emissionswinkel: $80^\circ \dots 89^\circ$) für eine Lufttemperaturzunahme um $1 \text{ K}/100 \text{ m}$.

Die Schallstrahlenverläufe ähneln rein qualitativ den Ergebnissen von Franck und Sager (1974), die ebenfalls mit linearen Temperaturprofilen Berechnungen durchgeführt haben. Einer Temperaturabnahme mit der Höhe entspricht die Ausbildung einer Schallschattenzone und damit Lärminderung bereits in geringer Entfernung von der Schallquelle. Eine Temperaturinversion demgegenüber begünstigt die Schallausbreitung.

Unter realen Bedingungen sind Temperatur- und Windeinfluss (Geschwindigkeit und Richtung) auf die Schallausbreitung jedoch immer gekoppelt zu betrachten. Eine einfache Beschreibung der Schallausbreitung mit **einem** Schallgeschwindigkeitsgradienten genügt daher für die Betrachtung bewegter Medien nicht.

3.3 Bestimmung der Schalldruckpegeldämpfung

Unter Verwendung der zweidimensionalen Schallstrahlensimulation wird eine entfernungsabhängige Dämpfung ermittelt (s. Raabe et al., 1998). Dazu werden von den Simulationsergebnissen die Emissions- und Immissionswinkel der Schallstrahlen (bezogen auf die Normale), die horizontale Entfernung der Auftreffpunkte der Schallstrahlen im Immissionsniveau und die Schallweglänge (Bogenlänge der Schallstrahlen) verwendet.

Die entfernungsabhängige Dämpfung wird auf eine Referenzintensität in der Entfernung x_{ref} von der Schallquelle bezogen. Dazu wird am Emissionsort eine Schallquelle der (dimensionslosen) Stärke Q angenommen. Die Referenzintensität ergibt sich dann unter Annahme einer ungestörten halbkugelförmigen sphärischen Ausbreitung des Schalls zu

$$\tilde{I}_{ref} = \frac{Q}{\pi x_{ref}} \text{ in m}^{-1} \quad (3.3)$$

Die Schallstrahlensimulation gibt die Möglichkeit, die einfache Annahme der kugelförmigen Ausbreitung der Schallwelle zu ersetzen. Die in dem durch zwei Schallstrahlen begrenzten Sektor emittierte Schallenergie erreicht nur Immissionsorte zwischen den Auftreffpunkten der Schallstrahlen im Immissionsniveau. Damit ist die räumliche Ausbreitung der Schallenergie so beschrieben, dass im Immissionsniveau die Schallenergie auf die Fläche senkrecht zur Ausbreitungsrichtung (also senkrecht zu den Schallstrahlen) verteilt wird. Die Schallintensität, ohne den Einfluss von Luftabsorption und Turbulenz, an einem Immissionsort zwischen zwei auftreffenden Schallstrahlen wird daher wie folgt berechnet:

$$\tilde{I}_{in} = \underbrace{\frac{Q|\alpha_{e1} - \alpha_{e2}|}{\pi}}_1 \cdot \underbrace{\frac{1}{|x_2 - x_1| \sin\left(\frac{\alpha_{i1} + \alpha_{i2}}{2} - \frac{\pi}{2}\right)}}_2 \text{ in m}^{-1} \quad (3.4)$$

Dabei gibt der erste Term die emittierte Schallenergie im Raumwinkel zwischen den zwei Schallstrahlen mit den Emissionswinkeln α_{e1} und α_{e2} wieder. Der zweite Term beinhaltet den Reziprokwert des zur den Strahlen normalen Abstandes zwischen den zwei Schallstrahlen am Immissionsort unter Verwendung des Mittelwertes der beiden Immissionswinkel α_{i1} und α_{i2} . Dabei sind x_1 und x_2 die horizontalen Entfernungen der Auftreffpunkte der Schallstrahlen im Immissionsniveau.

In einem einfachen numerischen Verfahren werden zunächst alle Schallstrahlen ausgewertet und die eingetragenen Schallenergien für äquidistante Entfernungsbereiche aufsummiert. Diese Schallenergien werden zur Referenzintensität in Beziehung gesetzt und daraus die Dämpfung für die Entfernungsbereiche ermittelt:

$$L_{D_I} = -10 \cdot \log \left(\frac{\tilde{I}_{in}}{\tilde{I}_{ref}} \right) \text{ in dB} \quad (3.5)$$

Eine Umrechnung der Intensitätsdämpfung in eine Schalldruckpegeldämpfung erfolgt mit

$$L_{D_p} = 2 \cdot L_{D_I} \text{ in dB.} \quad (3.6)$$

Bei den Berechnungen des Dämpfungspegels werden am Erdboden reflektierte Schallstrahlen nicht betrachtet, da hierfür die Bodeneigenschaften bekannt sein müssten, die Studie sich aber speziell dem Einfluss einer refraktierenden Atmosphäre auf die Schallausbreitung widmen sollte.

4. Meteorologisch bedingte Zusatzdämpfungen des Schalldruckpegels

4.1 Klasseneinteilung möglicher Schallausbreitungsbedingungen

Die meteorologische Beeinflussung der Schallausbreitung wird wesentlich durch die vertikalen Gradienten der Temperatur und des Windfeldes bestimmt. Dieser gekoppelte Einfluss der vertikalen Temperaturschichtung und der Windscherung kann durch ein Profil nicht ausreichend beschrieben werden. Es müssen vielmehr die Temperatur-, Windgeschwindigkeits- und Windrichtungsprofile getrennt betrachtet werden.

Um die Schallausbreitungsbedingungen für verschiedene meteorologische Situationen zu berechnen, müssen diese in verschiedene Klassen eingeteilt werden.

Dabei muss die große Vielzahl möglicher Stabilitäts- und Scherungsverhältnisse berücksichtigt werden. Die verwendete Höhenauflösung für die Temperatur-, Windgeschwindigkeits- und Windrichtungsprofile richtet sich hierbei nach vorgegebenen Daten (Modell- oder Messdaten), welche in einem Routineprogramm eingelesen und den standardisierten Profilklassen zugeordnet werden sollen.

Als Ergebnis verschiedener Sensitivitätsuntersuchungen wurde eine Klasseneinteilung für die Gradienten von Lufttemperatur, Windgeschwindigkeit und -richtung, wie in der Tabelle 2 dargestellt, vorgenommen. Die Kombination dieser verschiedenen Profile miteinander ergibt 1560 verschiedene Profilklassen. Die Höhenauflösung ist ebenfalls der Tabelle 2 zu entnehmen. Eine Einbeziehung der Vertikalprofile über einen Höhenbereich von 500 m hinaus hatte für das Schallimmissionsfeld in 2 m Höhe und in einem Umkreis von bis zu 15 km ausgehend von einer Schallquelle keine Auswirkungen auf die Schallausbreitung.

Die Tabelle für die Einteilung der Temperaturgradienten ist folgendermaßen zu lesen. Das Temperaturprofil Nr. 1 (siehe Abb. 3) ergibt sich, wenn die Temperatur in der 1. Schicht (0-50 m) zunimmt ($\Delta T/100 \text{ m} \geq 0$) und in der 2. Schicht (50-250 m) abnimmt, wobei diese Abnahme mindestens $0,25 \text{ K}/100 \text{ m}$ ($-0,25 > \Delta T/100 \text{ m}$) betragen muss (Fall 1). Das gleiche Ergebnis für die Schallausbreitung (Temperaturprofil Nr. 1, Fall 2) ergibt sich bei einer Temperaturabnahme in der 1. Schicht ($0 > \Delta T/100 \text{ m}$) und einer Temperaturabnahme oder Temperaturzunahme in der 2. Schicht, die kleiner als $0,5 \text{ K}/100 \text{ m}$ ist ($0,5 > \Delta T/100 \text{ m}$). Die zwei Zahlen in einer Spalte und einer Zeile beziehen sich also auf die beiden Bereichsgrenzen ($x > \Delta T/100 \text{ m}$ bzw. $\Delta T/100 \text{ m} \geq x$) für das Intervall des Temperaturgradienten. Das Symbol „-“ bedeutet, dass eine Bereichsgrenze nicht vorhanden ist, z.B. nur Zunahme (beliebiger Betrag des Gradienten) oder nur Abnahme (beliebiger Betrag des Gradienten) der Temperatur mit der Höhe vorgegeben wird.

Tab. 2a: Einteilung von Temperaturprofilen nach den vorherrschenden vertikalen Gradienten in verschiedenen Höhengschichten.

Δ TEMPERATUR [K]

Profil-Nr.	1. Schicht 0-50 m		2. Schicht 50-250 m		3. Schicht 250-500 m	
	$>\Delta T/100m \geq$		$>\Delta T/100m \geq$		$>\Delta T/100m \geq$	
1	-	0	-0,25	-	-	-
	0	-	0,5	-	-	-
2	-0,75	-1,25	-	0,5	-	-
	-1,75	-	-	0,5	-	-
3	0,25	-0,75	-	0,5	-	-
	-1,25	-1,75	-	0,5	-	-
4	-	0	0,25	-0,25	0,25	-
					-	0,25
5	-	0	-	0,25	0,25	-
					-	0,25

Tab. 2b: Einteilung von Windgeschwindigkeitsprofilen nach den vorherrschenden vertikalen Gradienten in verschiedenen Höhengschichten.

Δ WINDGESCHWINDIGKEIT [m/s]

Windgeschwindigkeit v_h am Erdboden	Profil-Nr.	$>$	$\Delta u/100m$	\geq
			(0-500 m, 2 BLM-Schichten)	
$v_h < 1$ m/s	1	0,25		-
	2	0,75		0,25
	3	1,25		0,75
$1 \leq v_h < 3$ m/s	4	1,25		-
	5	1,75		1,25
	6	-		1,75
$3 \leq v_h < 5$ m/s	7	2,25		-
	8	-		2,25
$5 \leq v_h < 7$ m/s	9	2,75		-
	10	3,25		2,75
	11	-		3,25
$7 \leq v_h$ m/s	12	3,25		-
	13	-		3,25

Das Windgeschwindigkeitsprofil Nr. 1 erhält man, wenn die Windgeschwindigkeit an der Erdoberfläche kleiner als 1 m/s ist und die Windgeschwindigkeitszunahme mit der Höhe kleiner als 0,25 m/s/100 m ist.

Bei diesen Profildefinitionen für die Windgeschwindigkeit wurden Fälle mit einer Windgeschwindigkeitszunahme mit der Höhe betrachtet, da diese Situation unter den meisten meteorologischen Bedingungen vorkommt. Die Windgeschwindigkeitszunahme wurde außerdem zu der am Boden vorhandenen Windgeschwindigkeit in Beziehung gesetzt. Höheren Windgeschwindigkeiten entsprechen in der Regel auch höhere Windgeschwindigkeitsgradienten.

Bei einer möglichen Abnahme der Windgeschwindigkeit mit der Höhe wird je nach der Windgeschwindigkeit am Boden das Profil Nr. 1, 4, 7, 9 oder 12 ausgegeben.

Die Profile für die Windrichtung wurden entsprechend der vorherrschenden Windrichtung am Erdboden und der Windrichtungsänderung in der 2. Schicht eingeteilt. Das Profil Nr. 1 ergibt sich bei einer Bodenwindrichtung von 345° bis 15°, d.h. Nordwind, sowie einer Änderung der Windrichtung in der 2. Schicht von weniger als 4°/100 m.

Dabei wurde von einer häufig vorkommenden Rechtsdrehung des Windes mit der Höhe ausgegangen. Bei einer Linksdrehung wird je nach Bodenwindrichtung das Profil Nr. 1, 3,... oder 23 ausgegeben.

Tab. 2c: Einteilung von Windrichtungsprofilen nach den vorherrschenden Bodenwindrichtungen und den vertikalen Gradienten in der 2. Höhengschicht.

Δ WINDRICHTUNG [Grad]

Profil-Nr.	Am Boden		2. Schicht (50-250 m)	
	> α (in 0m)	\geq	> $\Delta\alpha/100m$	\geq
1	15	345	4	-
2	15	345	-	4
3	45	15	4	-
4	45	15	-	4
5	75	45	4	-
6	75	45	-	4
7	105	75	4	-
8	105	75	-	4
9	135	105	4	-
10	135	105	-	4
11	165	135	4	-
12	165	135	-	4
13	195	165	4	-
14	195	165	-	4
15	225	195	4	-
16	225	195	-	4
17	255	225	4	-
18	255	225	-	4
19	285	255	4	-
20	285	255	-	4
21	315	285	4	-
22	315	285	-	4
23	345	315	4	-
24	345	315	-	4

In der Abbildung 3 sind Beispiele für entsprechende Profilklassen zusammengestellt.

Dabei wurden Temperaturprofile mit einer Temperaturabnahme (Nr. 1, Fall 2), mit einer freien Inversion ab 50 m Höhe (Nr. 3, Fall 2) und einer Bodeninversion (Nr. 5, Fall 1) gewählt. Die Windvektorprofile beinhalten eine Windgeschwindigkeitszunahme bis in 250 m Höhe sowie 2 Profile für die Windrichtung, wobei hier die Gradienten konstant sind und sich nur die Bodenwindrichtung (0° und 180°) ändert.

Mit der vorgestellten Einteilung von Temperatur- und Windvektorprofilen können alle möglichen Vertikalverteilungen dieser meteorologischen Größen erfasst werden. Es sind jedoch in extremen Situationen auch extreme Gradienten denkbar, welche die gewählten (äußeren) Intervallgrenzen um einen hohen Prozentsatz über- bzw. unterschreiten und damit auch zu einer signifikant geänderten Schalldämpfungskarte führen. Außerdem sind Spezialfälle des vertikalen

len Windvektorgradienten vorstellbar (z.B. Abnahme der Windgeschwindigkeit mit der Höhe), die z.Z. möglicherweise nur eine unzureichende Berücksichtigung in den Schalldämpfungskarten finden. Auf solche Fälle ist in einer praktischen Erprobungsphase besonders zu achten.

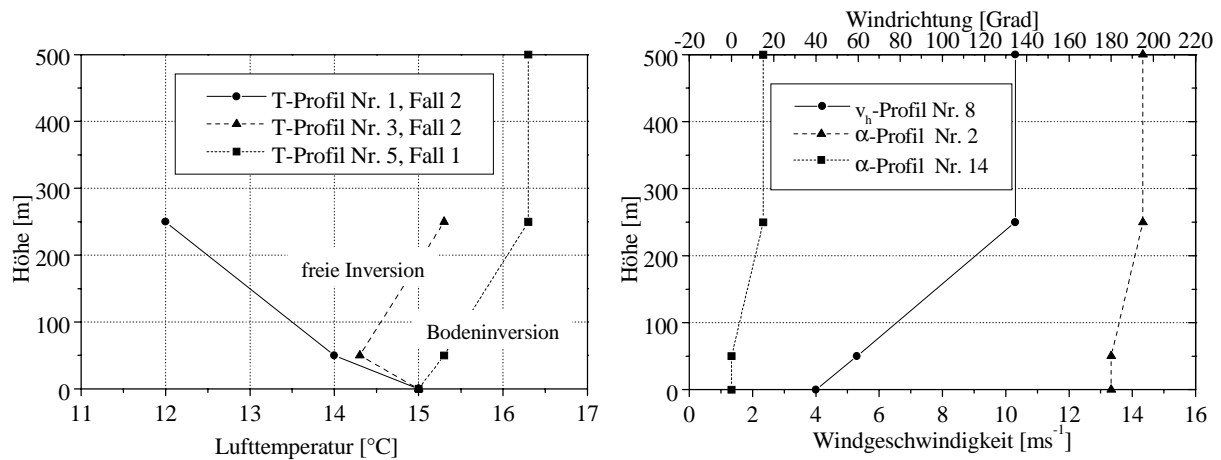


Abb. 3: Synthetische Vertikalprofile der Lufttemperatur T (links), der Windgeschwindigkeit v_h und Windrichtung α (rechts).

4.2 Auswertung

Die Berechnungen der Schalldruckpegeldämpfungskarten wurden mit dem LIM-Schallstrahlenmodell in einem Modellgebiet mit einem Abstand von 15 km um die Schallquelle durchgeführt. Das Untersuchungsgebiet wurde dabei mit einer horizontalen (radialen) Auflösung von 250 m und einer Winkelauflösung (Azimut) von 10° untersucht.

Als Emissionsniveau wurde die Erdoberfläche angenommen. Das Immissionsniveau liegt in einer Höhe von 2 m. Die Daten für dieses Höhenniveau wurden durch lineare Interpolation aus räumlich benachbarten Werten in 0 m und 50 m Höhe bereitgestellt.

Eingangsdaten waren die im Abschnitt 4.1 vorgestellten synthetischen Profile, wobei die verschiedenen Vertikalprofile für Lufttemperatur, Windgeschwindigkeit und -richtung miteinander kombiniert wurden.

Die Ausgabedaten sind Schalldruckpegeldämpfungswerte bezogen auf einen Referenzschalldruckpegel im Abstand von 1 m von der Schallquelle infolge der modifizierten Kugelwellendivergenz in Abhängigkeit von der Schallstrahlenrefraktion.

Ein Beispiel für die entsprechende Schalldämpfungsverteilung wird in Abbildung 4 gezeigt. Diese Dämpfungskarten zeigen ausschließlich den Beitrag der Schallstrahlenrefraktion an, d.h., es wurde die Differenz zwischen einer Dämpfungskarte mit Kugelwellendivergenz und Meteorologieeinfluss und einer Dämpfungskarte ausschließlich mit Kugelwellendivergenz ohne eine refraktierende Atmosphäre gebildet.

Werte um Null dB kennzeichnen den fehlenden Einfluss einer Schallstrahlenrefraktion, positive Werte stellen eine verstärkte Dämpfung (geringere Lärmbelastung durch eine Schallschattenzone) und negative Werte eine verringerte Dämpfung (verstärkte Lärmbelastung) des Schalldruckpegels infolge einer refraktierenden Atmosphäre dar.

Sehr hohe Werte kommen vor allem in den Schallschattenzonen zustande, in welche gemäß der Schallausbreitung nach der geometrischen Akustik kein Schallstrahl eindringt. Hier wurde ein Grenzwert für die Dämpfung (Kugelwellendivergenz+Meteorologieeinfluss) von 100 dB festgelegt, da ansonsten ein unendlich hoher Wert berechnet werden würde.

Die Dämpfungskarten zeigen in großen Gebieten eine starke Abweichung von der kugelsymmetrischen Verteilung bei einer Schallausbreitung ohne Meteorologieeinfluss.

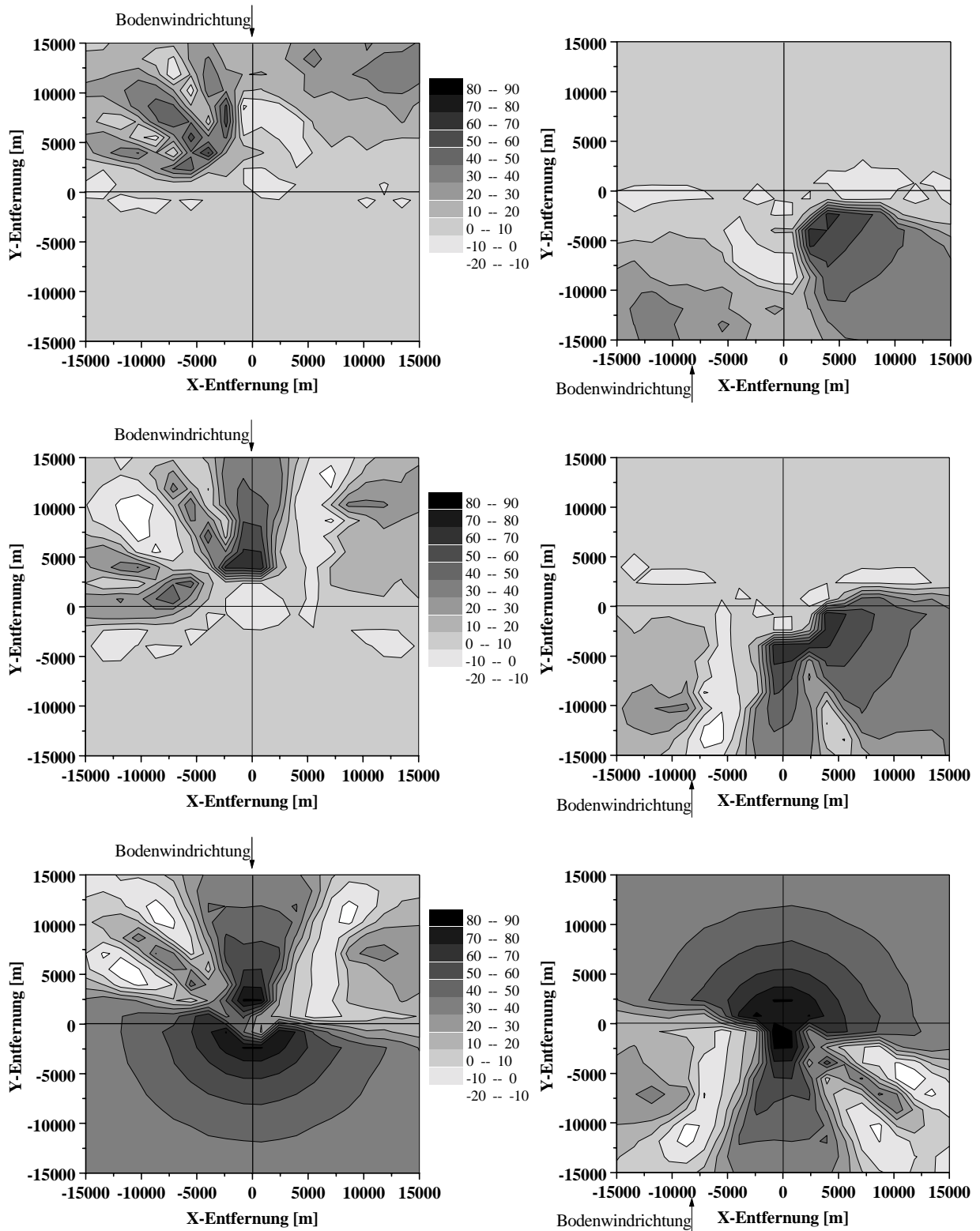


Abb. 4: Differenz aus einer Schalldruckpegeldämpfung [dB] mit und ohne refraktierende Atmosphäre für das Windgeschwindigkeitsprofil Nr. 8 (alle Bilder), das Temperaturprofil Nr. 1 (1. Zeile), 3 (2. Zeile) und 5 (3. Zeile) sowie das Windrichtungsprofil Nr. 2 (linke Spalte) und 14 (rechte Spalte). Eingangsdaten s. Abb. 1 und 2. Die Schallquelle befindet sich bei $(x, y)=(0, 0)$.

Bei Temperaturabnahme und bei einer freien Inversion (s. Abb. 4, 1. und 2. Zeile) entsteht eine Schallschattenzone in Gegenwindrichtung ausgehend von der Schallquelle, die besonders für das Temperaturprofil mit vertikaler Abnahme der Lufttemperatur ausgeprägt ist. Für den Fall einer Bodeninversion ergibt sich eine ausgedehnte Schallschattenzone in Mitwindrichtung von der Schallquelle.

Diese Ergebnisse geben den erwarteten Einfluss des Temperaturprofils auf die Schalldruckpegeldämpfung wieder.

Beim Vergleich der Dämpfungskarten für die unterschiedlichen Windrichtungsprofile (bei gleichem Windgeschwindigkeits- und Temperaturprofil) zeigt sich, dass trotz der erwarteten symmetrischen Verschiebung der Schallschattenzonen infolge der Drehung des Bodenwindes um 180° Unterschiede in einzelnen Regionen von bis zu 50 dB auftreten. In der Isolinien-darstellung werden diese Abweichungen der Abbildungen der linken und rechten Spalte nicht so deutlich, da das Plot-Programm besonders in der Umgebung starker Gradienten eine glättende Wirkung ausübt.

Anhand dieser Unterschiede wird deutlich, dass für eine Einschätzung der Schallausbreitungssituation in einem bestimmten Gebiet immer die aktuellen Kombinationen von Temperatur-, Windgeschwindigkeits- und Windrichtungsprofilen herangezogen werden müssen, da das Zusammenwirken dieser höhenveränderlichen Größen zu einem nicht pauschal abschätzbaren (z.B. mit Hilfe einer einzelnen Maßzahl oder eines einzelnen Profils) Ergebnis der Schalldruckpegeldämpfung führt.

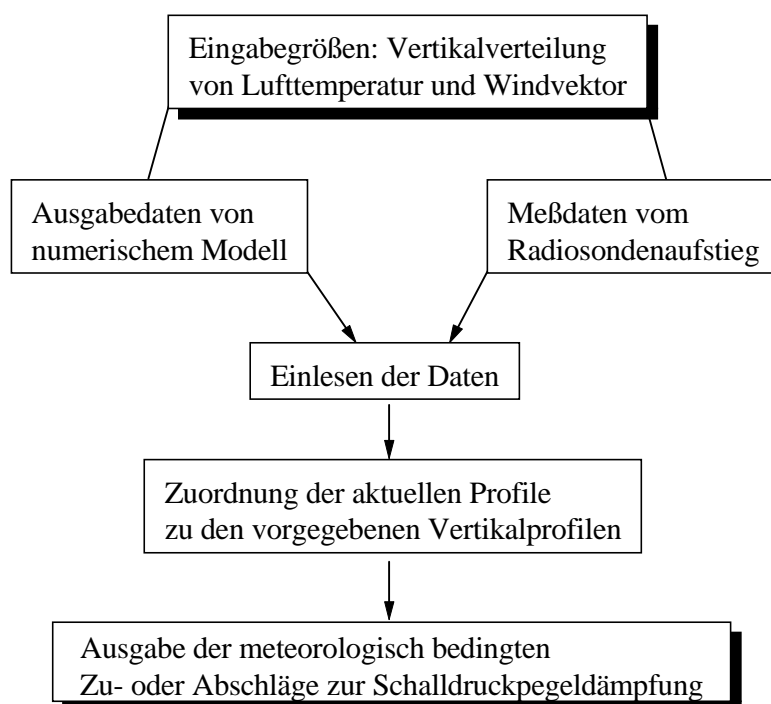


Abb. 5: Ablaufschema zur Ausgabe der gesuchten zusätzlichen Schalldruckpegeldämpfungswerte in Abhängigkeit von den Vertikalprofilen von Lufttemperatur, Windgeschwindigkeit und Windrichtung.

Die mit Hilfe dieses Verfahrens berechneten Dämpfungskarten werden nun in einer Datenbank zusammengefasst, welche für die standardisierten Profilkombinationen aus Temperatur-, Windgeschwindigkeits- und Windprofil die meteorologisch bedingte zusätzliche Schalldruckpegeldämpfung (positiv oder negativ) bereitstellt.

Um von einem aktuellen Profil, z.B. aus einem Atmosphärenmodell oder aus einer Radisondenmessung, auf die entsprechende Schallausbreitungssituation zu schließen, kann nach einem Ablaufschema gemäß Abbildung 5 vorgegangen werden, das mit Hilfe einer Programmumgebung verwirklicht wurde.

Hierbei werden die aktuellen Vertikalprofile der meteorologischen Größen eingelesen und aus diesen Informationen die passende standardisierte meteorologische Situation und die damit berechnete Dämpfungskarte herausgesucht.

Wenn eine klimatologische Datengrundlage für einen bestimmten Ort vorhanden ist, kann anhand der statistischen Auswertungen dieser Informationen festgestellt werden, wie häufig eine standardisierte Profilkategorie tatsächlich vorgekommen ist. Im Anschluss daran kann eine Schallausbreitungssituation als typisch oder untypisch für den ausgewählten Ort charakterisiert werden. Damit wird z.B. für die Beratung zu Lärmschutzfragen oder zur optimalen Beschallung unter Freiluftbedingungen ein Instrument bereitgestellt, das unter Einbeziehung der meteorologischen Bedingungen eine quantitative Einschätzung und objektive Bewertung der Schallausbreitungssituation zulässt.

5. Zusammenfassung und Schlussfolgerungen

Das Ziel dieser Studie bestand darin, ein Routineverfahren für die Lärmprognose unter Beachtung des Atmosphäreinflusses auf die Schallausbreitung zu entwickeln.

Die Auswertungen zeigen, dass bei einer bodennahen Schallquelle und einem Immissionsniveau in 2 m Höhe besonders die Profile von Lufttemperatur und Windvektor innerhalb der atmosphärischen Grenzschicht Beachtung finden und in möglichst hoher zeitlicher und räumlicher Auflösung als Eingangsdaten für die Schallausbreitungsmodellierung bestimmt werden müssen. Dabei sind die Gradienten von Lufttemperatur, Windgeschwindigkeit und -richtung von entscheidender Bedeutung für den Verlauf der Schallstrahlen. Die Bestimmung eines mittleren „Schallklimas“ für einen Ort, d.h., die Ermittlung von durchschnittlichen meteorologischen Bedingungen für die Schallausbreitung, führt infolge der durch den Mittelungsprozess stark abgeschwächten Vertikalgradienten der meteorologischen Größen zu einem Lärmkataster, das für den größten Teil der Einzelfälle (aktuelle Tagessituation) als wenig repräsentativ betrachtet werden muss.

Deshalb wurden vielfältige Schallausbreitungsbedingungen definiert sowie standardisiert und die zugehörigen Schallpegeldämpfungskarten in einer Datenbank erfasst. Diese Datenbank beinhaltet den meteorologischen Einfluss (Refraktion von Schallstrahlen) als Zu- bzw. Abschlag zu einem vorher berechneten Schalldruckpegel in Abhängigkeit von der Entfernung zur Schallquelle.

Mit dieser Datenbank und einem Programm, welches im Routinebetrieb einsetzbar ist, können Aussagen zur Bewertung von Lärmmessungen getroffen sowie ein Überblick zu den aktuell herrschenden Schallausbreitungsbedingungen gewonnen werden. Hierbei sollten die Resultate der Datenbank, die insbesondere den Beitrag meteorologisch bedingter Schallstrahlenrefraktion beinhalten, als Ergänzung zu vorhandenen Schallausbreitungsmodellen betrachtet werden. Hervorzuheben ist die nunmehr vorhandene Möglichkeit, mit dem bereitgestellten Routineverfahren eine objektive Bewertung einer vorliegenden Schallausbreitungssituation vorzunehmen. Dieses Resultat kann von keinem bisher genutzten Modell bzw. Berechnungsverfahren geliefert werden.

Es wird somit ein einfach zu bedienendes Instrument bereitgestellt, um die Lärmmessungen in der vom Gesetzgeber definierten Form bewerten zu können, ohne auf detaillierte Aussagen zu den Besonderheiten einer aktuellen Schallausbreitungssituation verzichten zu müssen.

Danksagung

Hiermit möchten wir uns für die Unterstützung durch R. Posselt und S. Hintschich während der Auswertung des umfangreichen Datenmaterials und der Vervollkommnung der numerischen Algorithmen bedanken. Außerdem danken wir M. Börngen für die kritische Durchsicht des Manuskripts.

Wir danken dem Bundesministerium für Verteidigung für die finanzielle Unterstützung (Auftragsnummern: E/E91S/Y0922/X5114 und E/E91S/Z0520/X5114) dieses Projektes.

Literatur

- Birkhoff, G., 1988: The consistency of models of sound waves in fluids. In: Lee, D. Sternberg, R.L. und Schultz, M.H. (Ed.), *Computational Acoustics. Wave propagation*, Elsevier Science Publishers B.V., 117-156.
- Bundes-Immissionsschutzgesetz (BImSchG), 1992: *Durchführungsvorschriften, Umweltrahmengesetz, UVP-Gesetz, Umwelthaftungsgesetz*, Dt. Fachschr.-Verl., Wiesbaden.
- Embleton, T.F.W., 1996: Tutorial on sound propagation outdoors. *J. Acoust. Soc. Am.* **100**, 31-48.
- Foss, R.N., 1979: Effects of wind and ground plane attenuation on sound propagation near the ground. *J. Acoust. Soc. Am.* **66**, 1088-1092.
- Franck, H. und Sager, G., 1974: Zur Schallwegverteilung in der unteren Grundsicht bei Vorhandensein einer Bodeninversion. *Z. Meteorol.* **24**, 130-137.
- Hallberg, B., Larsson, C. und Israelsson, S., 1985: Measurements of meteorological effects on longrange sound propagation using a m-sequence correlation. *J. Acoust. Soc. Am.* **78**, 1038-1044.
- Hirsch, K.-W. und Buchta, E., 1993: Zum Standardverfahren für die Berechnung der Schallimmissionen in der Umgebung von Truppenübungsplätzen. *Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 19. Deutschen Jahrestagung für Akustik/DAGA 93*, 656-659.
- Kriebel, A.R., 1971: Refraction and attenuation of sound by wind and thermal profiles over a ground plane, *J. Acoust. Soc. Am.* **51**, 19-23.
- Kühner, D., 1993: Schallausbreitung im klimatologischen Mittel. *Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 19. Deutschen Jahrestagung für Akustik/DAGA 93*, 299-308.
- L'Espérance, A.L., Nicolas, J., Wilson, D.K., Thomson, D.W., Gabillet, Y. und Daigle, G., 1993: Sound propagation in the atmospheric surface layer: Comparison of experiment with FFP predictions. *Appl. Acoustics* **40**, 325-346.
- Lamancusa, J.S. und Daroux, P.A., 1996: Ray tracing in a moving medium with two-dimensional sound-speed variation and application to sound propagation over terrain discontinuities. *J. Acoust. Soc. Am.* **93**, 1716-1726.
- Li, K.M., Taherzadeh, S. und Attenborough, K., 1998: An improved ray-tracing algorithm for predicting sound propagation outdoors. *J. Acoust. Soc. Am.* **104**, 2077-2083.
- Ostashev, V.E., 1997: *Acoustics in moving inhomogeneous media*. E & FN Spon, London, 259 S..
- Pierce, A.D., 1989: *Acoustics. An introduction to its physical principles and applications*, *Acoustic. Soc. Am.*, New York, 678 S..
- Piercy, J.E., Embleton, T.F.W. und Sutherland, L.C.: 1977, Review of noise propagation in the atmosphere. *J. Acoust. Soc. Am.* **61**, 1403-1418.

- Pridmore-Brown, D.C., 1962: Sound propagation in a temperature- and wind-stratified medium. *J. Acoust. Soc. Am.* **34**, 438-443.
- Raabe, A., Ziemann, A., Jagusch, F. und Arnold, K., 1998: *Studie: Berücksichtigung atmosphärischer Schallausbreitungsbedingungen bei Messungen und Prognose von Schießlärm*. Abschlußbericht zum Auftrag E/E91S/Y0922/X5114, 43 S..
- Raabe, A., Arnold, K. und Ziemann, A., 2000: *Berücksichtigung atmosphärischer Schallausbreitungsbedingungen bei Messungen und Prognose von Schießlärm*. Abschlußbericht zum Auftrag E/F360/X5144, 46 S..
- Sager, G., 1974: Eine Methode zur Bestimmung der Schallausbreitung in quasiruhender Atmosphäre. *Z. Meteorol.* **24**, 117- 129.
- TA Lärm, 1968: *Allgemeine Verwaltungsvorschrift über genehmigungsbedürftige Anlagen nach §16 der Gewerbeordnung - GewO. Technische Anleitung zum Schutz gegen Lärm*, 16.07.1968.
- TA Lärm, 1998: *Sechste Allgemeine Verwaltungsvorschrift zum Bundes-Immissionsschutzgesetz (Technische Anleitung zum Schutz gegen Lärm - TA Lärm)*, 26.08.1998.
- Turton, J. D., Bennetts, D. A. und Nazer, D. J. W., 1988a: The Larkhill noise assessment model, Part I: Theory and formulation. *Meteorol. Magazine* **117**, 145-154.
- Turton, J. D., Bennetts, D. A. und Nazer, D. J. W., 1988b: The Larkhill noise assessment model, Part II: Assessment and use; *Meteorol. Magazine* **117**, 169-179.
- VDI 2714, 1988: *Schallausbreitung im Freien*, Beuth Verlag, Berlin.
- Wiener, H.M. und Keast, D.N., 1959: Experimental study of the propagation of sound over ground. *J. Acoust. Soc. Am.* **31**, 724-733.
- Wöhle, W., 1984: Schallreflexion und Schallabsorption. In: Fasold, W., Kraak, W. und Schirmer, W. (Ed.): *Taschenbuch Akustik*. VEB Verlag Technik, Berlin, 55-82.
- Ziemann, A., 2000: Eine theoretische Studie zur akustischen Tomographie in der atmosphärischen Grenzschicht, *Wiss. Mitt. Inst. für Meteorol. Univ. Leipzig und Inst. für Troposphärenforschg. Leipzig* **19**, 141 S..

Anschrift der Autoren:

Astrid Ziemann (ziemann@rz.uni-leipzig.de), Klaus Arnold und Armin Raabe
 Institut für Meteorologie
 Universität Leipzig
 Stephanstr. 3
 04103 Leipzig

Acoustic tomography as a method to characterize measuring sites

A. Ziemann, K. Arnold and A. Raabe

Summary:

The method of acoustic tomography, based on external sonic energy, is applied inside the atmospheric surface layer to observe near-surface temperature fields. Important advantages of this technique as compared to other measurement methods are their remote-sensing capacity and the possibility to directly derivate area-averaged meteorological quantities.

The needed input data for the tomographically inverse algorithm are provided by the interaction of sound waves with the scanned atmospheric layer. The resulting horizontal slices lead to statements on the inhomogeneity of the underlying surface which may result in noticeable difficulties during the analysis of measuring campaigns with conventional methods.

Zusammenfassung:

Die auf der Aussendung von Schallenergie basierende Methode der akustischen Tomographie wird in der atmosphärischen Bodenschicht angewendet, um bodennahe Temperaturfelder zu beobachten. Bedeutende Vorteile dieses Verfahrens im Vergleich zu anderen Meßmethoden sind die Fernerkundungskapazität und die Möglichkeit, flächengemittelte Werte meteorologischer Größen direkt abzuleiten.

Die für den tomographischen Invertierungsalgorithmus benötigten Eingangsdaten werden durch die Wechselwirkung von Schallwellen mit der durchstrahlten Luftschicht bereitgestellt. Die resultierenden horizontalen Schnittbilder führen zu Darstellungen der Inhomogenität der Oberfläche. Letztere können beachtliche Schwierigkeiten während der Analyse von Messkampagnen mit konventionellen Methoden hervorrufen.

1. Introduction

Sound waves propagate through the atmosphere with different sound velocities according to the air temperature and wind vector fields influenced by the environmental conditions. Therefore, acoustic parameters characterising the sound propagation lead to a spatial description of the medium properties.

In the presented study measured travel time values of sound signals between different transmitters and receivers were used as initial line-integrated values to derivate spatially averaged quantities. Because of the information content of each measurement regarding to the properties of the atmosphere radiated through, a tomographic algorithm can provide a distribution of the meteorological quantities temperature and wind vector.

Such spatially averaged meteorological quantities are needed among others for the preparation and analyses of measuring campaigns as well as for the evaluation of model-output data, e.g. Large-Eddy simulations (Wyngaard and Peltier, 1996). By using the method of acoustic travel time tomography, it is, for instance, possible to derive information on the representativeness of point measurements and the homogeneity of measuring sites and therefore on the applicability of turbulence theories in data analysis and in atmospheric models.

The following section describes the method of acoustic tomography and its applicability to measurements inside the atmospheric boundary layer. In section 3, an overview of the experimental results from a measuring campaign at the "Common measuring field" at the Meteorological Observatory Lindenberg (Germany) is given. Thereby, the degree of inhomogeneity of

the measuring field will be described. The last section summarizes the results and supplies an outlook to future work.

2. Method of acoustic tomography

2.1. Application to the atmosphere

As mentioned above there is a great demand for spatially averaged meteorological data to directly validate models and to describe the characteristics of measuring sites. The needed data could be provided on one side by point measurements and a following interpolation procedure or on the other side by remote-sensing methods which yield data from a volume, an area or along a line. Acoustic travel time tomography with a specially experimental equipment and distribution of measuring devices as well as a special analysis algorithm belongs to the last group of methods (see Fig. 1).

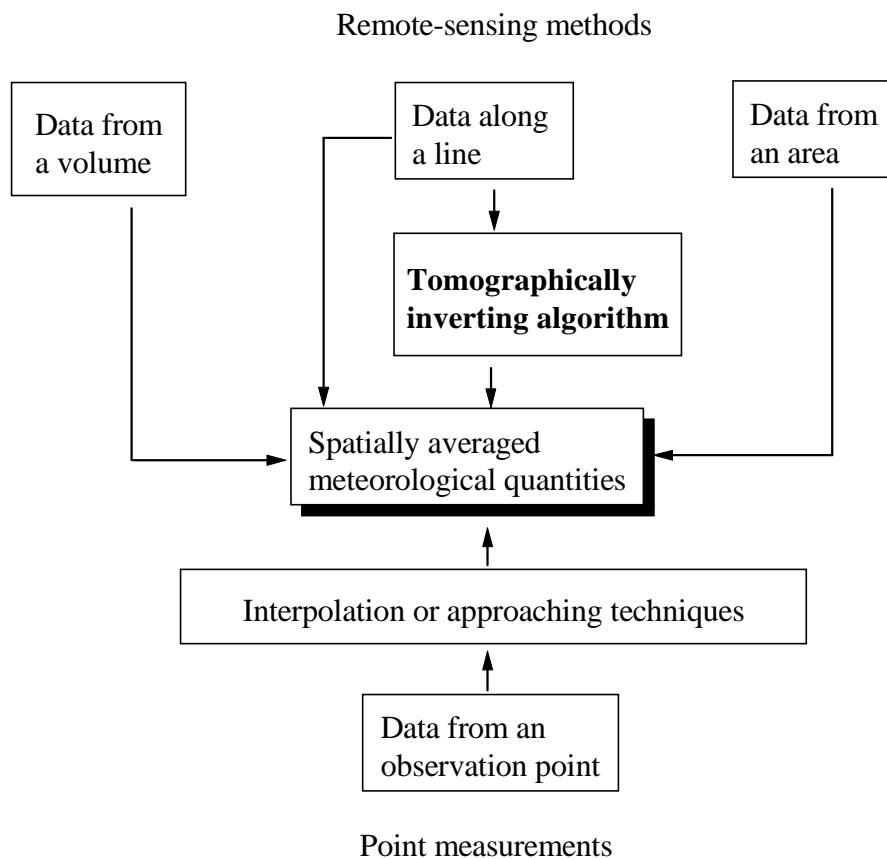


Figure 1: Possibilities to derive spatially averaged meteorological quantities.

Thereby, the acoustic travel time tomography is mainly characterised by the high accuracy regarding to other remote-sensing methods and by the creation of horizontal slices through the atmosphere (Ziemann, 2000).

In our study tomography is defined as a combined measurement and data analysis technique to reconstruct a slice through a medium using the reaction of the considered medium on an external probing energy. For the special case of acoustic travel time tomography measurements have to be performed on different ray paths between the acoustic sources and receivers. The

rays must touch all parts of the experimental area, best with a quite uniform distribution. The derivation of spatially averaged values of physical parameters, e.g. effective sound speed including the influences of air temperature and wind vector, results from the inverting of all single travel time measurements.

The procedure describes the opposite situation in comparison to the traditional forward problem. Thereby, by outgoing from values of measured quantities, one can use mathematical inverting techniques to derive an estimation of the values of system parameters that explain or reproduce the experimental observations.

The possible application of acoustic travel time tomography to the atmospheric surface layer was validated by Spiesberger and Fristrup (1990) as well as Wilson and Thomson (1994). In contrast to these studies the application of an other tomographic algorithm will demonstrate how one can detect absolute values of meteorological quantities without additional information (see Ziemann et al., 1999b; Ziemann, 2000).

2.2. Simultaneous Iterative Reconstruction Technique

The measured quantity, that means the line integral of the acoustic travel time of a signal between a fixed transmitter and a receiver, can be expressed as

$$\tau = \int_{\text{ray}} \frac{dl}{c_{\text{eff}}} \quad (1)$$

where dl is the element of the arc length along the propagation path and

$$c_{\text{eff}} = v + c_L(T_{\text{av}}) \quad (2)$$

symbolizes the effective sound speed with the coupled influence of the virtual acoustic temperature T_{av} (c_L is Laplace's sound speed) and the wind vector component in direction of the sound propagation v (e.g., Spiesberger and Fristrup, 1990).

Because of this relation it is possible to derive meteorological quantities like the air temperature by means of acoustic travel time tomography under the prerequisite of a high measurement accuracy which depends on the searched meteorological quantity, the sound-receiver distance and the environmental conditions.

The coupled influence of the virtual acoustic temperature and the wind vector on the effective sound speed is a difficulty in the data analysis. One possibility to distinguish between these different influences and to derive a travel time without the wind vector influence is described by Arnold et al. (2001). Thereby, one quantity (e.g., wind vector) is iteratively changed whereas the variance of the other quantity (e.g., temperature) decrease stepwise for all ray paths together until reaching a designated criterion. In the next step this procedure will be carried on with the reversed quantity. At last one obtains two data sets: one travel time set only with the temperature influence and one travel time set only with the wind vector influence.

An additional difficulty in the numerical application of the tomographic method is the dependency of the ray path itself on the unknown distribution of the effective sound speed and therefore the line integral becomes non-linear in this quantity. Usually, a linearization is applied to solve this difficulty (e.g., Bording et al., 1987). Thereby, straight lines connecting the sound source and the receiver are used to approximate the true ray path. The error made by this approximation was investigated by sensitivity tests with a sound-ray model including a generalised equation for refraction. The error found will be small enough if we use path

lengths not more than 500 m over a relatively homogeneous surface (see Ziemann et al., 1999b; Ziemann, 2000).

As a result of the limited number of transmitters and receivers the tomographic algorithm has to use a tomographic array divided into grid cells of finite size.

After the linearization and discretization as well as the removing of wind influence from the travel time the system of equations, which has to be solved for i sound rays and j grid cells, results as

$$\tau_i = \sum_{j=1}^J l_{ij} s_j \quad (3)$$

with the so called slowness s_j (reciprocal sound velocity c_{Lj}). In the case of travel time tomography, one has to solve the inverse problem to get the spatially averaged sound speed depending on the virtual acoustic temperature. By outgoing from values of the measured travel time, an inverting technique provides an estimate of the spatially resolved temperature field that reproduces the measurements.

There are various algorithms to solve the linear equation system, for instance iterative algebraic reconstruction techniques which were developed for applications in seismic tomography (Gilbert, 1972) and for medical purposes (Herman, 1980).

Especially the Simultaneous Iterative Reconstruction Technique (SIRT) is characterised by stable convergence during the application on data sets with measuring errors, by non-significant developments of artefacts as well as simple handling during online evaluation (see Humphreys and Clayton, 1988; Ziemann et al., 1999a).

A comprehensive overview of the mathematical background of SIRT and other iterative algebraic reconstruction techniques is given by van der Sluis and van der Vorst (1987). These inverting algorithms follow a similar scheme.

An initial guess of the slowness values in the grid cells is derived from a simple back projection of the measured travel time data into grid cells using the inverse Eq. 3. By means of the so estimated reciprocal sound speed one can apply forward modelling (Eq. 3) to get a modelled travel time. Now a difference between the experimentally obtained and the simulated travel time values is provided by the model. After back projection of this difference and adding a resulting correction to the present model an updated version of the simulated travel time follows. This iterative improvement of the modelled data is continued until attainment of a convergence criterion and leads to area-averaged values of temperature T_{av} .

Using the straight-line sound rays and the described SIRT algorithm allows to derivate absolute values of the temperature without additional information in contrast to stochastic reconstruction algorithms which were used by Spiesberger and Fristrup (1990) or Wilson and Thomson (1994) in the atmospheric surface layer.

2.3. Accuracy

The travel time measurements as well as the determination of the distances between sources and receivers have to be carried out with a high degree of accuracy. Additionally, the achieved data accuracy principally depends on the examined meteorological quantity, the transmitter-microphone distance and the environmental conditions (Arnold et al., 1999; Ziemann et al., 1999b). Furthermore, the quality of the tomographic reconstruction has to be checked for the special tomographic array with synthetic data. The results of such tests of the SIRT model demonstrate the relative insensitivity of the algorithm with regard to small variations of the transmitter and receiver positions up to ± 5 cm (Ziemann, 2000).

With an actual travel time accuracy of about 0.3 milliseconds the temperature field can be obtained with an indefiniteness of 0.5 K (spatial average) for one measurement.

3. Experimental results and discussion

3.1. Data analysis

The spatial resolution of a measuring field and therewith the grid cell size in the SIRT model mainly depends on the number of sound rays between transmitters and receivers, the accuracy of the travel time measurements and the differences between the sound speed values at different places inside the measuring area.

For the tomographic measuring site at the "Common measuring field" (German Weather Service Observatory Lindenberg) 70 km south-east of Berlin with a horizontal dimension of $200 \times 240 \text{ m}^2$ we could apply a grid cell size of $50 \times 50 \text{ m}^2$ apart from the borders (see Fig. 2).

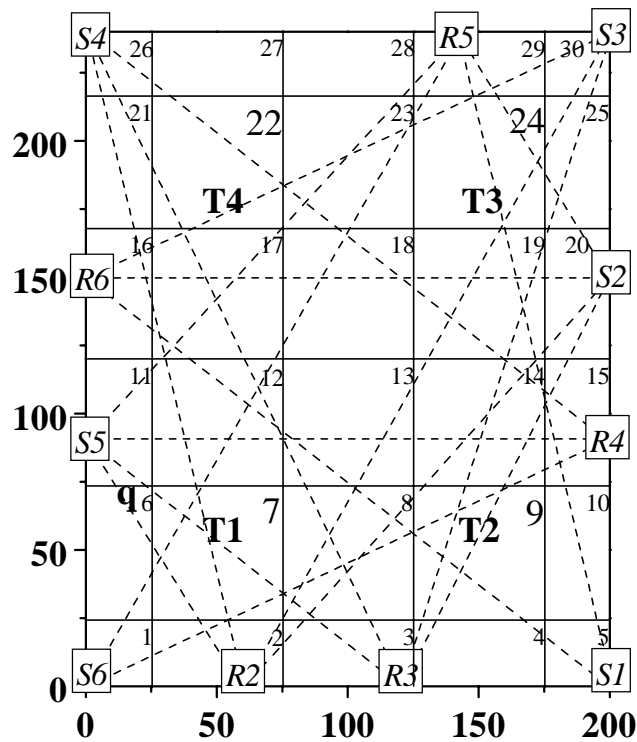


Figure 2: Tomographic array ($200 \times 240 \text{ m}^2$) with 6 sources (S1...S6), 5 receivers (R2...R6), 4 thermo-electric measurement devices (T1...T4), 1 measurement point for the specific air humidity (q), numbered grid cells (1...30) and sound rays (dashed lines).

By means of 6 compression drivers at a height of about 2 m using a 1000 Hz signal and 5 microphones at the same height, the travel times between the sources and receivers are determined.

Details of the experiment and the technical devices are described by Arnold et al. (2001).

To derive the air temperature depending on T_{av} the specific humidity has to be included in the calculations. For this purpose humidity measurements (one point) of the German Weather Service Lindenberg are used.

3.2. Spatially averaged values of temperature

With the adjusted travel time data (removed wind influence) area averages of the air temperature at a height of 2 m were calculated with one value for one grid cell.

Figure 3 illustrates examples of so-called tomograms, that are horizontal slices of the air temperature field at the site.

Beside the general temperature trend, for instance, the daily course of the air temperature, one can detect more or less significantly spatial differences of the air temperature (Fig. 3, left). Such pictures can lead to statements on the horizontal variability of the temperature field and additionally to statistical data inspections also to statements on the horizontal homogeneity of measuring fields.

A tomogram (Fig. 3, right) with averaged air temperatures over a time period of 24 hours demonstrates the relative homogeneity of the measuring field in relation to temperature measurements.

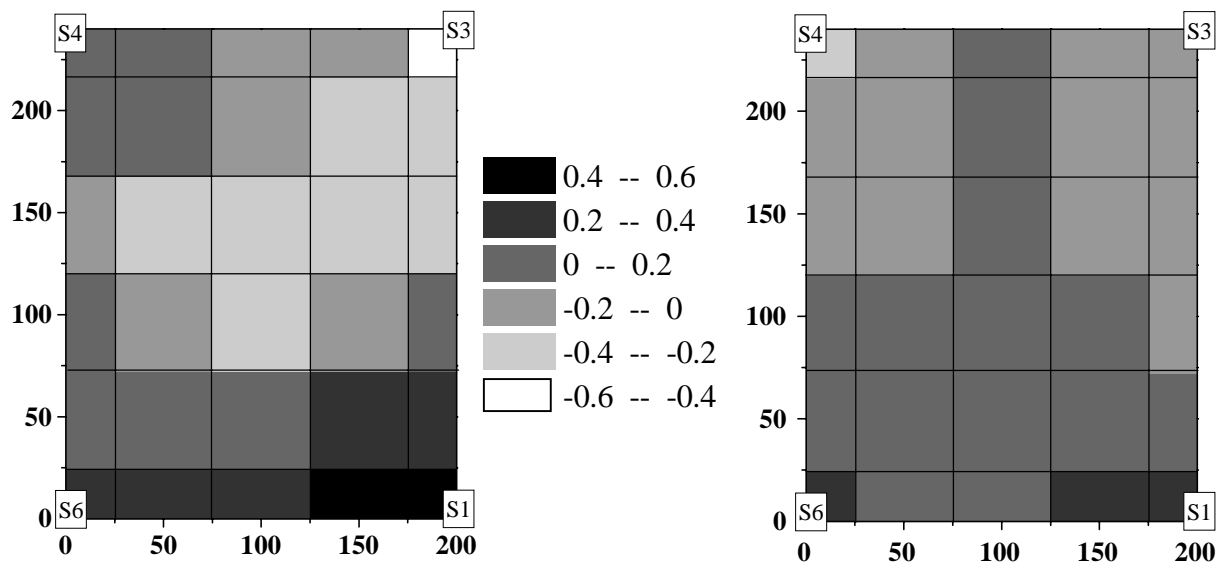


Figure 3: Tomograms of the air temperature differences between the grid cell value and the areally averaged value [K].

Left: 10-Minutes-average of the air temperature on the 24 September 1999, 1200 UTC.

Right: 24-hours-average of the air temperature on the 24 September 1999.

Figure 4 shows the comparison between two spatial differences of the air temperature measured by the thermocouples and derived by tomography (see also Fig. 2). Sometimes one can determine a qualitatively good agreement between the two methods, that means the same sign of the difference. This result especially occurs during weak wind conditions (horizontal wind speed $< 2 \text{ ms}^{-1}$). Sometimes one has to notice a disagreement in the signs of the differences, especially in connection with higher wind speeds. The wind influence on the two measuring methods, for instance, on the effective sound speed measured by tomography as well as on advection, is one possible reason for this behaviour. Further possible explanations for the differences between the tomographic and the conventionally derived data are:

- Differences between remote-sensing method and direct measurement
- Differences between area-averaged data and data from one point
- Non-homogeneous humidity field (tomography)
- Different influence of local advection

In spite of these differences between the two methods one can detect the general agreement in the absolute values of the temperature differences and in the spatial variability of these differ-

ences. A higher variability of the spatial differences and also greater absolute values of these differences appear during the day between 0700 and 1300 UTC with a developed turbulence under convective conditions. This development was interrupted by a thunderstorm with a strong shower. The acoustic measurements has to be stopped during this period because the signal-to-noise ratio is too bad in this situation (noise of the drops on the microphone) and the microphones are not really waterproof.

Please note, that the measurement cycle during the day amounts to 30 s, whereas the measurements during the night were repeated every 5 minutes. This is in addition to the more distinct turbulence a reason for the obviously greater variability of the spatial temperature differences during the day. Furthermore, the mapped values are instantaneous values without averaging.

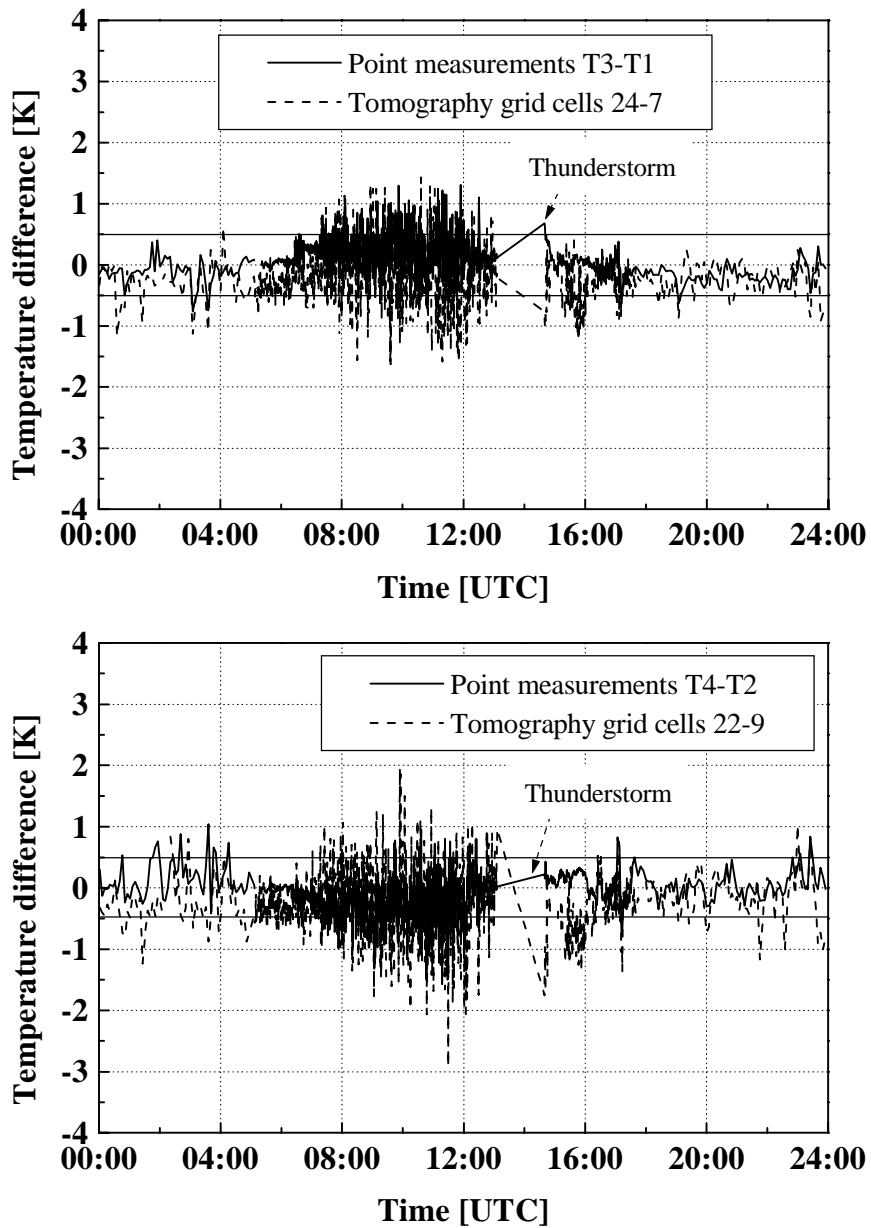


Figure 4: Spatial temperature differences (point-to-point and cell-to-cell, respectively) measured by thermocouples (solid line) and by tomography (dashed line) on the 24 September 1999) every 30 seconds during the day (05:06-17:39 UTC) and every 5 minutes during the night.

In Figure 4 the indefiniteness of the absolute temperature values measured by tomography is also indicated. However, the differences of values measured by one and the same method can reach a higher accuracy.

The calculation of statistical measures over the chosen time period is a further possibility to describe the homogeneity of the measuring field.

Therefore, the standard deviation of grid cells in relation to a spatially averaged value for the whole tomogram (see Fig. 5) and the normalised third moment (skewness, see Fig. 6) were calculated for the selected time period.

The standard deviation achieves for the most grid cells only a small value between 0.2 K and 0.4 K. That means the deviation of the single value inside the grid cell from the spatially averaged value is smaller than the indefiniteness of the measurement for one time step (ca. 0.5 K). Furthermore there are only small variations between different grid cells. Therefore, the measuring field can be indicated as horizontally homogeneous for this time period.

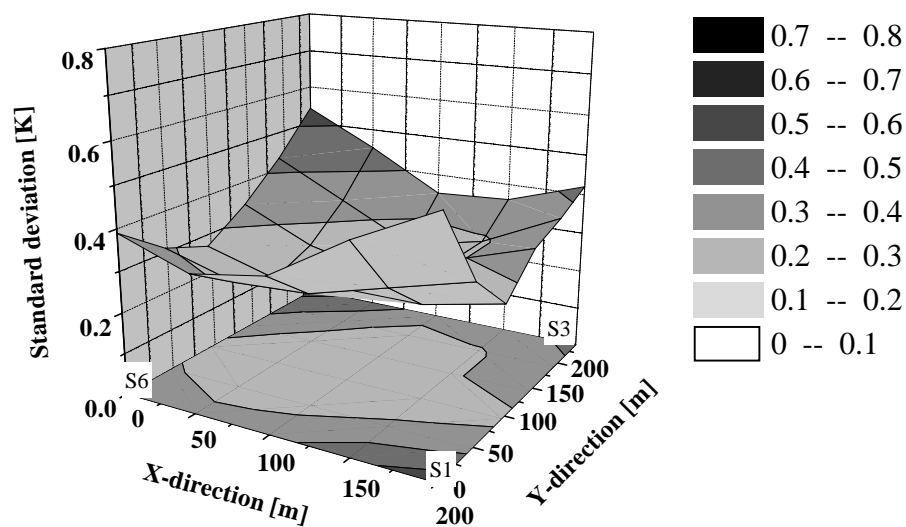


Figure 5: Standard deviation of the tomographically derived air temperature inside the grid cells in relation to a spatially averaged value, 0000-2400 UTC, 24 September 1999.

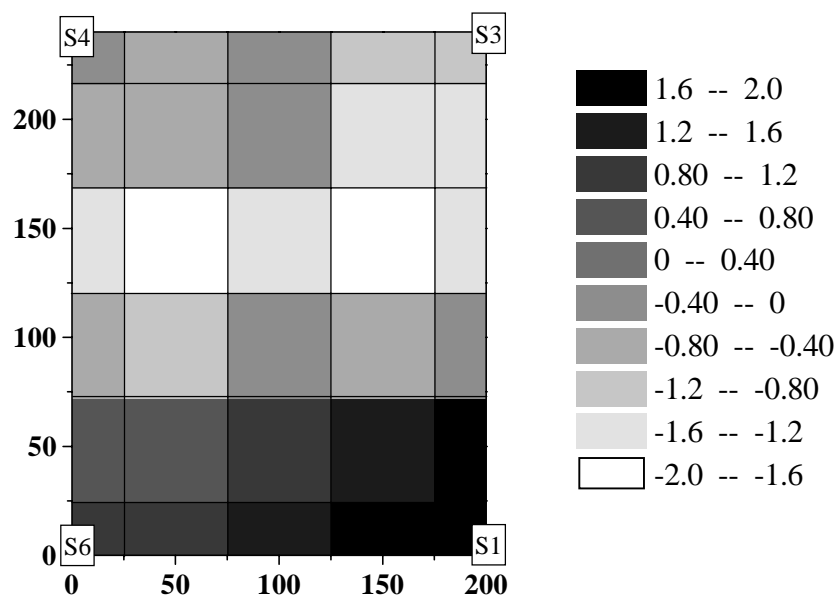


Figure 6: Momentum coefficient of skewness of the tomographically derived air temperature inside the grid cells in relation to a spatially averaged value, 0000-2400 UTC, 24 September 1999.

In comparison to this results the normalized third moment displays a spatially greater variability. This measure denotes the type of the data distribution, i.e. whether the distribution is symmetrical or not. Positive (negative) values signify that the most frequent value is smaller (greater) as the arithmetic mean. A really symmetric distribution has none skewness.

The significant spatial differences of this quantity (significance test: Madansky, 1988) point to the horizontal inhomogeneity of the measurement field. This could be a hint for different environmental conditions and turbulent processes that influence the different parts of the test site. Further investigations with other analysis methods (e.g., Fourier analysis) and by means of additional measurements are necessary to identify and quantify these processes.

4. Conclusions and outlook

The acoustic travel time tomography holds the possibility to derive statements on the variability of spatially averaged air temperature data and therefore on the possible application of the known turbulence theories to calculate vertical turbulent fluxes.

Except for the differences between point measurements and spatially averaged data the qualitatively sufficient agreement between the two methods point out the horizontal homogeneity of the measuring field in relation to the air temperature at least over time periods of several hours.

Therefore point measurements could be applied for micro-meteorological investigations over the examined area of the “Common measuring field” under convective conditions during a late summer day with the known restrictions regarding to the disadvantages of direct measurements (e.g., influence of direct radiation to absolute temperature measurements) and with sufficiently extensive averaging time intervals in mind.

The degree of inhomogeneity of a measuring field should be tomographically analysed however under different environmental conditions (e.g., meteorological situation, type of the underlying surface) to get a generalised statement on the representativeness of point measurements and the applicability of usually used turbulence theories and approximations (e.g., horizontal homogeneity and stationarity) for the calculation of turbulent fluxes.

In the future the validation of the tomographic data has to be carried on. Besides, a comparison with results from other tomographic methods will be provided to increase the quality and reliability of the reconstructed data.

Further investigations will deal with a refined separation of the wind-temperature influence on the effective sound speed to calculate also tomograms of the wind vector.

Additionally, three-dimensional tomographic measurements and algorithms will be developed to provide volume-averaged meteorological data. By using these data it should be possible to develop improved objective measures of the degree of inhomogeneity and to experimentally test such measures which were discussed in the meteorological community (see, e.g., Panin et al., 1998; Zilitinkevich and Calanca, 2000). Such investigations could be an important step to evaluate and quantify the effects of micro-scale inhomogeneity of natural surfaces on the flux measurements at one point and the determination of the energy balance of an area.

Acknowledgments

We would like to thank F. Weiße and M. Engelhorn as well as S. Schienbein for their support in the development and manufacturing of the tomographic system and the fast response thermocouples, respectively. Furthermore, we acknowledge the staff of the German Weather Service Lindenberg, especially F. Beyrich, as well as K. Radtke and H. Heinrich from the LIM for their assistance during the observations. We also wish to thank N. Mölders for helpful comments and suggestions.

References

- Arnold K., A. Ziemann and A. Raabe, 1999: Acoustic tomography inside the atmospheric boundary layer. *Phys. Chem. Earth (B)*, **24**, 133-137.
- Arnold K., A. Ziemann and A. Raabe, 2001: Acoustic tomography in comparison to in-situ temperature and wind measurements. *Wiss. Mitt. Inst. f. Meteorol. Univ. Leipzig*, **22**, (this issue).
- Bording, R.P., A. Gersztenkorn, L.R. Lines, J.A. Scales and S. Treitel, 1987: Applications of seismic travel-time tomography. *Geophys. J. Roy. Astr. Soc.*, **90**, 285-303.
- Gilbert, P.F.C., 1972: Iterative methods for three-dimensional reconstruction from projections. *J. Theor. Biol.*, **36**, 105-117.
- Herman, G.T., 1980: *Image reconstruction from projection*. Academic Press, Orlando, 1-283.
- Humphreys, E. and R.W. Clayton, 1988: Adaption of back projection tomography to seismic travel time problems. *J. Geophys. Res.*, **93**, 1073-1085.
- Madansky, A., 1988: *Prescriptions for working statisticians*. Springer, New York, 295 pp.
- Panin, G.N., G. Tetzlaff and A. Raabe, 1998: Inhomogeneity of the land surface and parameterization problems of surface fluxes in natural conditions. *Theoret. and Appl. Climatol.*, **60**, 163-178.
- Spiesberger J.L. and K.M. Fristrup, 1990: Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography. *Am. Natural.*, **135**, 107-153.
- Van der Sluis, A. and H.A. van der Vorst, 1987: Numerical solution of large, sparse linear algebraic systems arising from tomographic problems. *Seismic tomography: With applications in global seismology and exploration geophysics*. G.D. Nolet, Ed., Reidel Publishing Company, Dordrecht, 49-83.
- Wilson D.K. and D.W. Thomson, 1994: Acoustic tomographic monitoring of the atmospheric surface layer. *J. Atm. Ocean. Technol.*, **11**, 751-768.
- Wyngaard, J.C. and L.J. Peltier, 1996: Experimental micrometeorology in an era of turbulence simulation. *Boundary-Layer Meteorol.*, **78**, 71-86.
- Ziemann, A., K. Arnold and A. Raabe, 1999a: Acoustic tomography in the atmospheric surface layer. *Ann. Geophysicae*, **17**, 139-148.
- Ziemann A., K. Arnold and A. Raabe, 1999b: Acoustic travel time tomography. A method for remote sensing of the atmospheric surface layer. *Meteorol. Atmosph. Phys.*, **71**, 43-51.
- Ziemann, A., 2000: Eine theoretische Studie zur akustischen Tomographie in der atmosphärischen Grenzschicht. *Wiss. Mitt. Inst. f. Meteorol. Univ. Leipzig u. Inst. f. Troposphärenforschg. Leipzig*, **19**, 141 pp. (thesis in German)
- Zilitinkevich, S. and P. Calanca, 2000: An extended similarity theory for the stably stratified atmospheric surface layer. *Q. J. R. Meteorol. Soc.*, **126**, 1913-1923.

Address:

Astrid Ziemann (ziemann@rz.uni-leipzig.de), Klaus Arnold und Armin Raabe
Institut für Meteorologie
Universität Leipzig
Stephanstr. 3
04103 Leipzig

Acoustic tomography in comparison to in-situ temperature and wind measurements

K. Arnold, A. Ziemann and A. Raabe

Summary:

Acoustic travel time tomography is presented as an experimental technique for remote monitoring of areally averaged meteorological quantities as the air temperature and the horizontal wind speed.

This ground based remote sensing technique uses the nearly horizontal propagation of sound waves in the atmospheric surface layer. Here the acoustic travel time tomography was applied by measuring the travel time at defined propagation paths between several sound sources and receivers. The resulting sound speed were used to obtain estimates of the meteorological parameters.

A measuring campaign was carried out at the test site in Lindenberg (DWD) to compare the acoustically derived data with conventional systems. These observations demonstrated that on one side the accuracy of the acoustic system is comparable with in-situ measurements and on the other side the temperature was particularly significant overestimated by the standard sensors, e.g. due to the radiation influence.

Zusammenfassung:

Die Akustische Laufzeittomographie wird als ein experimentelles Verfahren zur Sondierung meteorologischer Parameter, wie z.B. der Lufttemperatur und der horizontalen Windgeschwindigkeit, vorgestellt.

Dieses bodengebundene Fernerkundungsverfahren nutzt die horizontale Ausbreitung von Schallwellen in der atmosphärischen Grenzschicht. Hier wird das Verfahren der Laufzeittomographie angewendet, d.h. bei bekannter Weglänge wird die Ausbreitungszeit von ausgesendeten Schallsignalen zwischen mehreren Schallquellen und Empfängern gemessen. Die resultierenden Schallgeschwindigkeitsinformationen werden genutzt, um daraus die entsprechenden meteorologischen Parameter abzuleiten.

Auf dem Gelände des Meteorologischen Observatoriums Lindenberg (DWD) wurde eine Messkampagne durchgeführt, um die akustischen Sondierungen mit konventionellen Systemen zu vergleichen. Die Auswertungen zeigen, dass einerseits die Genauigkeit der Akustischen Tomographie vergleichbar mit den konventionellen in-situ Messungen ist und andererseits, dass die Lufttemperatur aufgrund des Strahlungseinflusses bei Messungen mit den üblichen Sensoren zum Teil erheblich überschätzt wird.

1. Introduction

Spatially averaged data derived from ground based measurements were conventionally provided by point measurements and additional interpolation algorithms. A relatively new way to get such values directly and with a highly spatial and temporal resolution is the transfer of tomographic methods to the atmospheric surface layer.

There are various advantages of tomographic measurements compared with conventional methods, first the effect as a spatial filter for turbulence below the micro-scale, second the remote monitoring, because the test medium is not influenced by devices, and third a higher

number of data per sensor in comparison to the traditional point measurements is obtained (Wilson & Thomson, 1994).

As an inverse technique tomography has been used routinely for instance in medicine or geophysics for many years. The successful application of acoustic tomography schemes for a fluid, e.g. the ocean, was executed for instance by Clay & Medwin (1977) and Munk & Wunsch (1978). The transfer of the acoustic tomography to the atmosphere is more critical because under these conditions, an immutable state of the investigation area during the measurement, can't be presumed.

However Spiesberger & Fristrup (1990) as well as Wilson & Thomson (1994) demonstrated the applicability of acoustic tomography to the atmospheric boundary layer. The significant difference to our current work lies in the numerical algorithm and the experimental realisation. In contrast to the cited publications in this study area-averaged absolute values of the temperature and wind speed are estimated without additional measurements.

Thereby, a type of acoustic travel time tomography was used, where the sound speed can be determined by measuring the travel time of a signal at a defined propagation path. Applying a suitable procedure, measurements of the speed of sound can be used to reconstruct the spatially and temporally variable temperature and wind fields (Ziemann et al., 1999b, Raabe, et al., 2001).

2. Sound Speed in the Atmosphere

The sound propagation can be considered as spreading of small density perturbations in an unbounded, frictionless and compressible fluid. If air is treated as an ideal gas and under adiabatic conditions, the Laplace equation of sound speed can be used:

$$c_L = \sqrt{\gamma R T_{av}} \quad (1)$$

where c_L is the Laplace sound speed, γ is the specific heat ratio, R is the gaseous constant of dry air and T_{av} is the acoustic virtual temperature. The temperature derived from the speed of sound in air is similarly to the virtual temperature – defined as the temperature at which dry air has the same density as moist air at the same pressure – however the variable specific heat of water vapour and dry air are considered:

$$T_{av} \cong T(1 + 0.513q) \quad (2)$$

where q is the specific humidity. Additionally to spatial and temporal alterations concerning changes of air temperature, also the influence due to the wind field appends to the sound velocity and leads to the (horizontal) effective sound speed:

$$\mathbf{c}_{eff}(T_{av}, \mathbf{v}) = c_L(T_{av}(t, x, y))\mathbf{n} + \mathbf{v}(t, x, y). \quad (3)$$

So the measured sound speed depends on the air temperature, the humidity as well as the wind vector. To separate the different influences several methods are applicable, for instance to use –analogously to a sonic anemometer– the reciprocal sound propagation (Kaimal & Gaynor, 1991) or to solve numerically a system of equations (Ziemann et al., 1999a; Arnold, 2000).

In this study, for the separation of the influence of the temperature and the wind field on the sound speed, a special iteration loop was used, where the wind speed and direction was changed as long as the mean value deviation of the Laplace sound speed achieves a minimum:

$$\frac{1}{n} \sum c_L - \bar{c}_L \rightarrow \min. \quad (4)$$

Therefore, except for the humidity (which is assumed as horizontally homogeneous), no additional information is necessary and for each measurement the temperature and wind vector can be estimated. The iteration can be performed for the whole array or for different parts of the measuring area. Thus for each source–receiver connexion one value for the temperature and for the respective area the information about the wind speed and direction are available.

3. Experimental Realisation

To obtain meteorological relevant data the sound speed determination has to be carried out with a high degree of accuracy. This supposition includes on the one side the exact travel time measurements and on the other side the precise determination of the distances between sources and receivers as well as the knowledge of the actual sound paths.

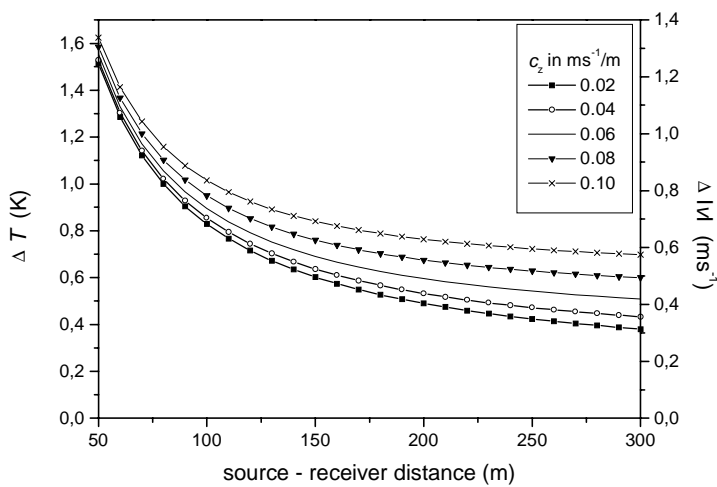


Fig. 1: Inaccuracy of the temperature (left) and wind speed (right) measurements due to the influence of path length (± 0.10 cm) and travel time measurements (± 0.3 ms) as well as the vertical effective sound speed gradient c_z depending on the source receiver distance.

Figure 1 shows the coupled influence of the measuring error, that contains the sound path length (± 0.10 cm) and the travel time (± 0.3 ms), and the actual meteorological conditions which were represented by the vertical effective sound speed gradient on the temperature and wind speed. With increasing sound speed gradient the ray paths are stronger curved and the deviation from the linear path increases likewise. Figure 1 demonstrates that for an adequate accuracy of the temperature (± 0.5 K) and wind speed (± 0.5 ms⁻¹) measurements a source receiver distance of more than 150 m is necessary.

A field experiment was carried out at the test site of the German Weather Service (DWD) in Lindenberg (south east from Berlin) in autumn 1999. Six sources and five receivers were positioned at the boarder of an array of 200×240 m². Figure 2 illustrates the layout of the tomographic array. The sources, compression drivers, and the receivers, microphones, were positioned on tripods at a height of 2 m, respectively. To compare the acoustically derived data with in-situ measurements inside the array 4 thermocouples (T1 ...T4), a sonic anemometer (both at 2 m above ground) and a 12 m wind- and temperature mast are used. The thermocouples are identically fast response sensors which were mounted in ventilated Young-shelters (Schienbein & Arnold, 1996).

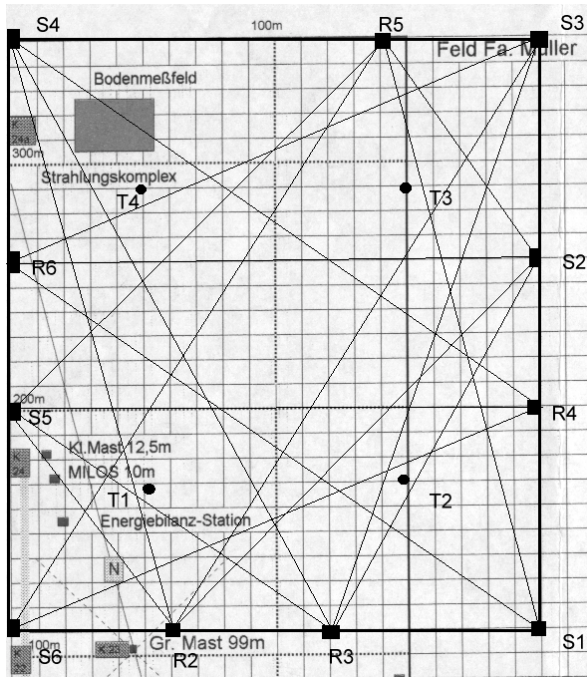


Fig. 2: Layout of the tomographic array at the DWD test side nearby Lindenberg ($200 \times 240 \text{ m}^2$). The sources are labelled S1 to S6, the receivers R2 to R6 (R1 is the frequency generator). T1 to T4 are the thermocouples; the sonic anemometer is positioned at the “Energiebilanz – Station” (the plotted grid cells are $10 \times 10 \text{ m}^2$).

All sources simultaneously transmit a short sine burst with a constant frequency of 1000 Hz (double peak of twice 4 ms). The assignment of each signal to the corresponding source is possible due to the tomographic array layout where each ray has a significantly different path length (see Fig. 3).

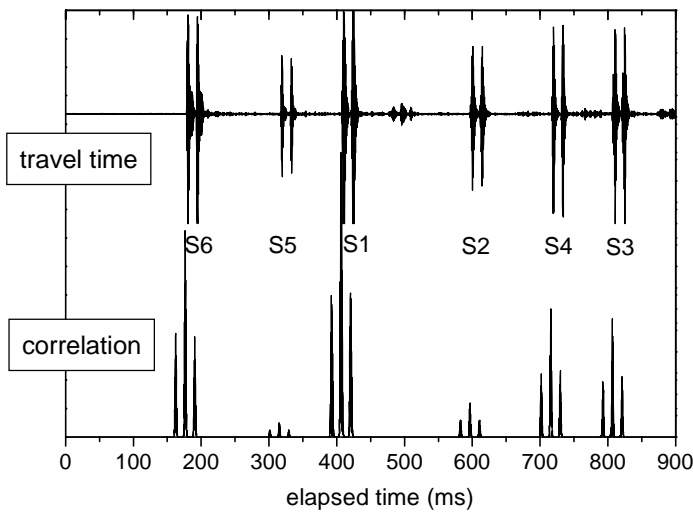


Fig. 3: Example of the recorded signals (Receiver R2) after bandpass filtering (top) and of the correlation coefficient with the reference signal (down) during the tomography experiment (18:51 UTC on 23 September 1999).

The travel time of each signal was estimated by special correlation algorithm between the received and the transmitted signal (Arnold, 2000). Each peak of the cross correlation is associated with a separate ray path. In Figure 3 the six arrivals with the appending correlation peaks according to the source number are visible. The correlation maximum corresponds with the travel time whereas the small deceleration between the signal start and the correlation peak is due the inertia of the acoustic system (Arnold et al., 1999).

4. Results

After separating the two prime effects on the effective sound speed, the temperature and wind speed can be calculated for each transmission path. Now the acoustically estimated data were compared with the conventionally derived meteorological measurements, which were located inside the tomographic array. As a result area-averaged data were compared with different point measurements.

Figure 4 and 5 show the comparison between the over all single paths averaged recalculated temperature and wind speed with the in-situ measurements. This comparison was carried out for a few days in autumn 1999 under nearly horizontal homogeneous conditions.

For the air temperature estimation (instead of the acoustic virtual temperature) an additional information on the humidity was used under the assumption that this measurement is representative for the whole array.

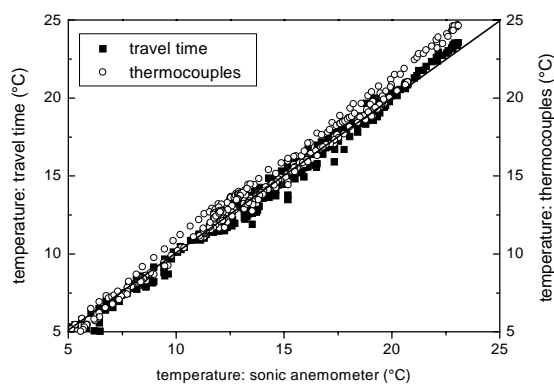


Fig. 4: Comparison between the conventional temperature measurements (sonic anemometer and thermocouples) and the acoustic derived temperature (arithmetic mean over all sound paths) on the 23/24 September and 12/13 October 1999. All values were averaged over 10 min.

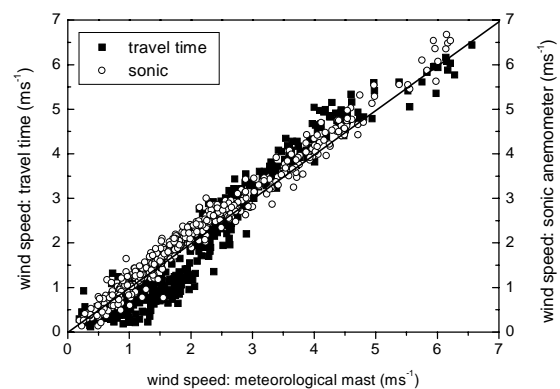


Fig. 5: Comparison between the conventional wind speed measurements (cup anemometer and sonic anemometer) and the acoustic derived wind speed (arithmetic mean over all sound paths) on the 23/24 September and 12/13 October 1999. All values were averaged over 10 min.

Figure 4 demonstrates a good agreement of the temperature for the different in-situ measurements as well as for the acoustically derived data. Especially the sonic anemometer data accord with the travel time data very well. However in Figure 4 a remarkable difference between the sonic (or travel time) measurements and the thermocouples is visible. Particularly at high values the temperature is overestimated by the thermocouples. This is primarily caused by the influence of radiation and the sensor inertia. For the wind speed (Fig. 5) a high analogy is identifiable too. However for wind velocities up to 2 ms^{-1} an underestimation of the travel time data is visible, probably due to the influence of different wind directions inside the array.

In Figure 6 the temporal course of air temperature for on day in September 1999 is plotted and the different measuring systems were compared. Here the sound speed data were recalculated into temperature values using a horizontal constant humidity and subsequent averaged over the 30 paths.

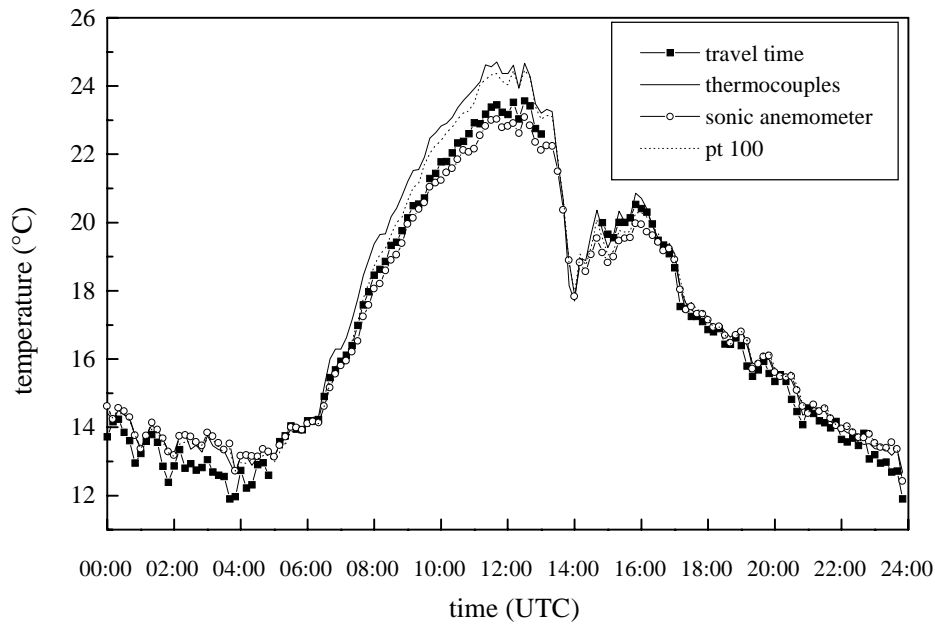


Fig. 6: Daily course of the air temperature registered by the different systems: travel time (arithmetic mean value over all sound paths), thermocouples (arithmetic mean value over 4 sensors), pt 100 (meteorological mast) and sonic anemometer on the 24 September 1999 at the test side Lindenberg (all values were averaged over 10 min). In the afternoon the measurements were intermittent by a thunderstorm.

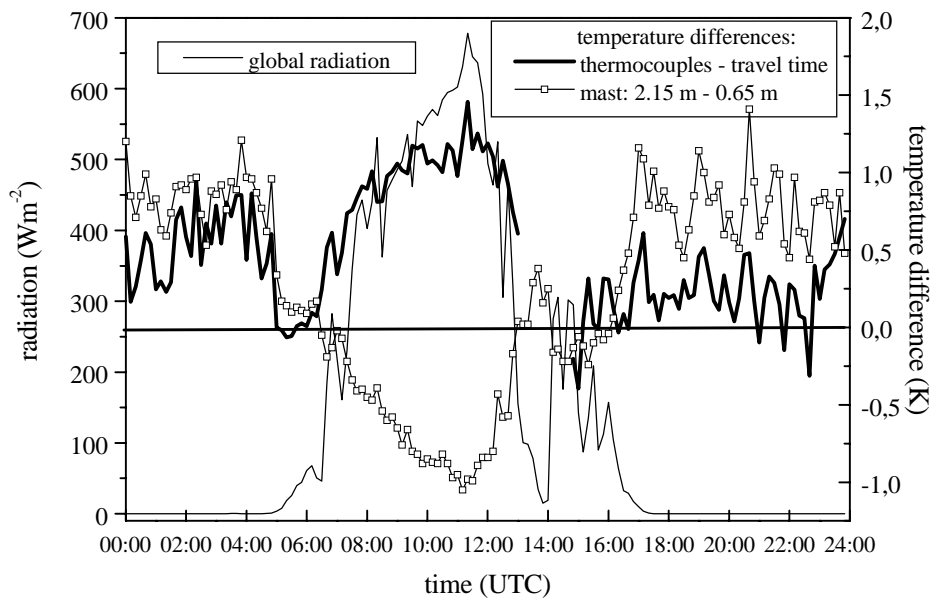


Fig. 7: Daily course of the global radiation (left) and the difference between the thermocouples and the travel time temperature (right) as well as difference of the temperature between 2.15 m and 0.65 m above the ground (on 24 September 1999 at the test side Lindenberg). All values were averaged over 10 min.

Due to the spatial dimensioning the acoustic data are representatively for an array of $200 \times 240 \text{ m}^2$ and thermocouples an array of $100 \times 120 \text{ m}^2$, whereas the pt 100 at the mast and the sonic thermometer/anemometer represent point measurements.

A high congruence between all independent measuring methods is cognizably in Figure 6. However at night and during the forenoon differences between the temperature measurements are discernibly.

During the cooling off phase at night the averaged travel time temperatures are from time to time about 1 K below all other measurements (see Fig. 7). This may be caused by a regional cooling at a boarder of the array, which was not detected by the point measurements, or as a result of the slope of the terrain were the sound paths crosses partially deeper layers. Figure 7 demonstrates that particularity for stronger vertical temperature gradients close to the ground (range between 2.15 and 0.65 m) the differences between the thermocouples and the sonic temperature increase.

However the differences during the heating period (from 06:00 till 13:00) are still stronger distinctive. Here the temperature measurements carried out by the thermocouples and the Pt 100 sensors are up to 1 K higher as the sonic anemometer/thermometer as well as the acoustic tomographic measurements (see Fig. 7). The reasons therefore are multilaterally: at first the radiation warming of the shelter, secondly the heat storage inside the shelter and at least the release of heat by the ventilation motor. Figure 7 demonstrates that after the sunrise with increasing solar radiation the differences between the thermocouples and the sonic temperature grow up in the same manner.

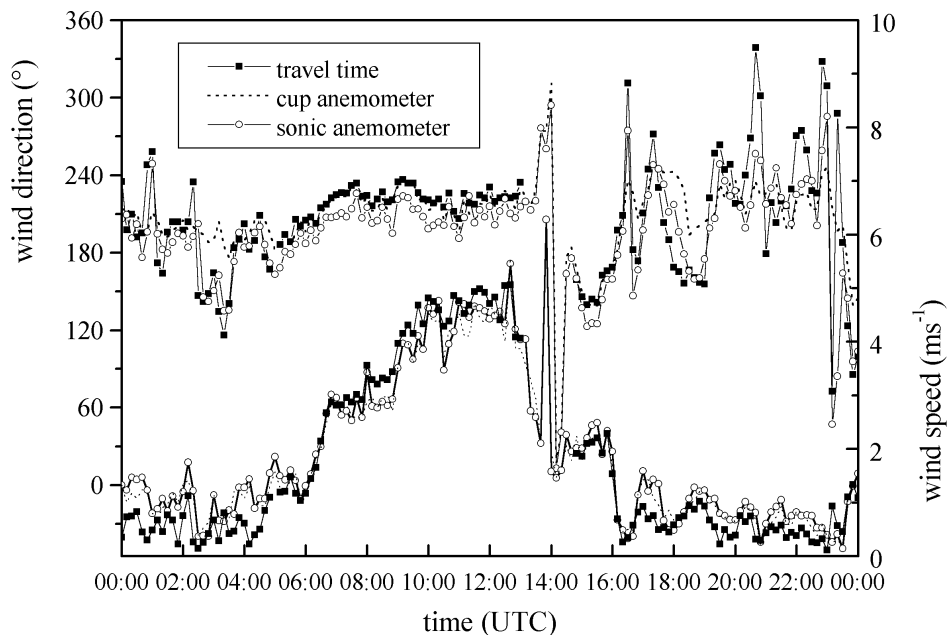


Fig. 8: Daily course of the wind speed (bottom range) and wind direction (top range) registered by different measuring systems: travel time (arithmetic mean over all sound paths), cup anemometer (mast) and sonic anemometer on the 24 September 1999. All values were averaged over 10 min.

Furthermore in Figure 8 the different wind velocity and direction measurements were compared. As already displayed in Figure 5, all three independent systems measure nearly the same wind speed. Particularly the wind direction agrees very well for all three independent measuring systems. Merely at low velocities the wind speed was underestimated by the acoustic system. These differences are probably due to the incomplete separation of the different in-

fluences on the sound speed as well as horizontally different wind velocity and direction inside the tomographic array.

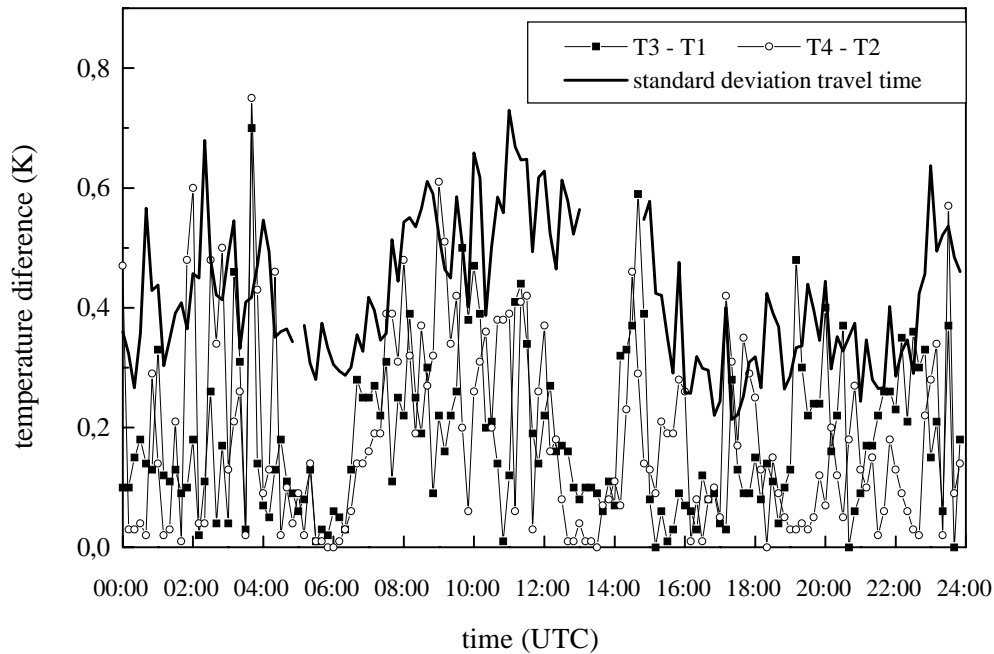


Fig. 9: Daily course of the absolute differences between the temperature sensors T3 - T1 and T4 - T2 (thermocouples) in comparison with the standard deviation of the acoustic derived temperature on the 24 September 1999. All values were averaged over 10 min.

Besides the area averaged values for the temperature and wind vector the spatial temperature differences are considered in Figure 9. Here the absolute differences between the temperature sensors were compared with the range of variation of the acoustic temperature measurements. Figure 9 demonstrates that even under nearly homogeneous conditions the over 10 minutes averaged temperature difference between two thermocouples reaches up to 0.7 K. Particularly at the forenoon this temperature inhomogeneities are conserved over longer periods. The standard deviation for the temperature of the 30 acoustic paths shows nearly the same behaviour. For time periods with stronger temperature differences between the thermocouples the standard deviation increases in the same manner.

5. Conclusions

The evaluation demonstrates the applicability of the tomographic monitoring of near surface horizontal temperature and wind fields. Temporal and spatial variances in the wind and temperature field are clearly veritable. The accuracy of the travel time as well as the sound path determination are sufficient to estimate the meteorological parameters with the adequate exactness.

One fundamental problem is the separation of the different effects on the effective sound speed. Here the humidity is assumed to be horizontally homogeneous and the temperature and wind influence was separated by an repetitive technique. By this means for each propagation path one value of the temperature and for the whole tomographic array the wind speed and direction can be calculated. This are the raw data for the following tomographic inversion.

In principle all different measuring systems record similarly values for the temperature as well as the wind speed and direction. Nevertheless it exists a noteworthy overestimation of the

temperature by the thermocouples and the platinum probe (pt 100). These differences are mainly due to the heat storage inside the sensor shelters, especially during stronger solar radiation. Because of all standard temperature measurements were carried out inside sensor shelters this fact is especially remarkable.

The tomographic monitoring is applicable, e.g., for the validation of micro-scale (LES) models with a resolution up to 100 m. Other uses are the determination of horizontal gradients of meteorological quantities or the detection of turbulent structures inside a convective boundary layer.

Acknowledgements

We would like to thank F. Weiße and M. Engelhorn for their support in the development and manufacturing of the acoustic measuring system. Special thanks are going to Dr. S. Schienbein for the development and maintenance of the thermocouples. Furthermore, we acknowledge Dr. F. Beyrich and the staff of the test site GM Lindenberg for the data providing and the cooperation during the field experiment. Likewise we would like to thank K. Radtke and H. Heinrich for their assistance during the experiment.

References

- Arnold, K., Ziemann, A. & Raabe, A., 1999: Acoustic tomography inside the atmospheric boundary layer. *Phys. Chem. Earth (B)*, **24**, 133-137.
- Arnold, K., 2000: Ein experimentelles Verfahren zur Akustischen Tomographie im Bereich der atmosphärischen Grenzschicht. *Wiss. Mit. Inst. f. Meteorol. Uni. Leipzig*, **18**.
- Clay, C.S. and Medwin, H., 1977: *Acoustical Oceanography. Principles and applications*. J. Wiley & Sons Inc., New York, 544 p.
- Kaimal, J.C. and Gaynor, J.E., 1991: Another look at sonic thermometry. *Boundary-Layer Meteorol.*, **56**, 401-410.
- Munk, W. and Wunsch, 1978: Ocean acoustic tomography: A scheme for large scale monitoring. *Deep-Sea Res.*, **26A**, 123-161.
- Wilson, D.K. and Thomson, D.W., 1994: Acoustic tomographic monitoring of the atmospheric surface layer. *J. Atm. Ocean. Techn.*, **11**, 751-768.
- Raabe, A., Arnold, K. and Ziemann, A., 2001: Near surface averaged air temperature and wind speed determined by acoustic travel time tomography. *Meteorol. Z.*, in print.
- Schienbein, S. and Arnold, K., 1996: Konzeption und Aufbau einer Meßeinrichtung für die Erfassung von Temperaturen und Strömungsgeschwindigkeiten bis 12 m Höhe. *Wiss. Mit. Inst. f. Meteorol. Uni. Leipzig*, **4**, 124-130.
- Spiesberger, J.L. and Fristrup, K.M., 1990: Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography. *Am. Natural.*, **135**, 107-153.
- Ziemann, A., Arnold, K. and Raabe, A., 1999a: Acoustic travel time tomography - a method for remote sensing of the atmospheric surface layer. *Meteorol. Atmos. Phys.*, **71**, 43-51.
- Ziemann, A., Arnold, K. and Raabe, A., 1999b: Acoustic tomography in the atmospheric boundary layer. *Ann. Geophysicae*, **17**, 139-148.

Address of authors:

Klaus Arnold, Astrid Ziemann and Armin Raabe
Universität Leipzig, Institut für Meteorologie
Stephanstr. 3, D-04013 Leipzig, Germany
(arnold@rz.uni-leipzig.de)

Messung der Lufttemperatur und Luftfeuchte unter Berücksichtigung des Strahlungsfehlers auf einem Versuchsgelände

S. Schienbein und K. Arnold

Zusammenfassung:

Zur Validierung des Verfahrens der akustischen Tomographie, das als Fernerkundungsverfahren trägheits- und berührungslos arbeitet, werden Vergleichsmessungen mit herkömmlichen Sensoren für die Lufttemperatur und Luftfeuchte durchgeführt. Dabei zeigen sich die Grenzen der Erfassung dieser Größen in Bezug auf Strahlungseinfluss und Trägheit der Messgeräte. Hier werden einige Probleme bei der Verwendung konventioneller Hütten und Lösungsvorschläge aufgezeigt.

Summary:

In order to validate measurements with the acoustic tomography, which operates as a remote sensing system lagless and contactless, comparison measurements were performed with conventional sensors for the air temperature and humidity. The accuracy of these quantities is limited by the radiation influence and the lag of the instruments. Here some difficulties are pointed out by using conventional shelters and solutions are demonstrated.

1. Einleitung

Das am Leipziger Institut für Meteorologie entwickelte Verfahren der akustischen Tomographie nutzt die Laufzeit von ausgesendeten Schallsignalen zwischen mehreren Sendern und Empfängern aus, um daraus meteorologische Parameter wie die Lufttemperatur oder das Windfeld zu bestimmen. Die Schallsender und -empfänger sind im Gelände so verteilt, dass aus der tomographischen Invertierung flächenhafte Verteilungen der entsprechenden Größen gewonnen werden können. Zur Validierung dieser Daten werden parallele Messungen mit konventionellen Sensoren, die mit möglichst hoher Genauigkeit und trägheitsarm die entsprechenden Parameter bereitstellen, benötigt.

Seit dem Betreiben von Messnetzen benutzen Meteorologen für die Messung von Lufttemperatur und -feuchte Wetterhütten, die baugleich gestaltet sind und durch eine spezielle Konstruktion, eine von der direkten Sonnenstrahlung unbeeinflusste Messung der Lufttemperatur gewährleisten sollen. In den letzten Jahrzehnten sind mit den neuen Sensoren auch eine Vielzahl von kleineren Hütten entstanden, die für die Messung der strahlungsgeschützten Temperatur eingesetzt werden. Besonders wenn unterschiedliche Messverfahren in Anwendung kommen, macht sich der immer noch unzulängliche Strahlungsschutz und die Trägheit des Messfühlers bemerkbar. Da die akustische Tomographie als ein indirektes Fernerkundungsverfahren die virtuelle Lufttemperatur trägheits- und berührungslos bestimmt, ist es für einen Vergleich mit konventionellen Sensoren besonders wichtig, die Trägheit und den Strahlungsschutz zu optimieren.

Will man Temperaturänderungen auf dem Versuchsgelände mit einer Ausdehnung von $300 \times 300 \text{ m}^2$ bezüglich der lokalen Unterschiede und der zeitlichen Veränderung betrachten, müssen die Temperatursensoren eine Genauigkeit von $\pm 0,1 \text{ K}$ besitzen und zusätzlich zum Strahlungsschutz auch rasch reagieren, um Schwankungen innerhalb einer Minute noch anzuzeigen. In den nachfolgenden Ausführungen soll über einen Vergleich von Temperaturmessun-

gen in einer Young-Hütte und einer Stevenson-Hütte, über das Trägheitsverhalten und Temperaturmessfehler in einer ventilierten Young-Hütte und über die Erprobung einer neuen ventilierten Hütte berichtet werden.

2. Vergleich von Temperaturlaufzeichnungen in der Young-Hütte und in der Stevensonhütte

Am Institut für Meteorologie der Universität Leipzig laufen bereits mehrere Jahre Aufzeichnungen der verschiedensten meteorologischen Messgrößen. Insbesondere werden Minutenmittelwerte der Lufttemperatur in einer Young-Hütte und in einer Stevenson-Hütte in 2 m über dem Boden und etwa 3 m horizontal entfernt aufgezeichnet. Diese Aufzeichnungen wurden über einen Zeitraum von 2 Jahren verglichen und ergaben folgende Ergebnisse:

Im Vergleich der einzelnen Minutenmittelwerte finden wir zum Teil erhebliche Abweichungen. Im Tagesmittel aller 1440 Messwerte sind die Unterschiede sehr gering (unter der Fehlergrenze von 0,1 K). Diese Erkenntnis ergibt sich aus der Auswertung aller Differenzen von Minutenmittelwerten aus der Young-Hütte und der Stevenson-Hütte über den Tag. Von jedem Tag wurden die größten negativen und positiven Abweichungen ermittelt und dann noch über den Monat gemittelt. In der Tabelle 1 sind die über den Monat gemittelten Abweichungen der Tagesmittel, der größten negativen Abweichung während eines Tages von dem Messwert der Young-Hütte und der größten positiven Abweichungen eines Tages von dem Messwert der Young-Hütte eingetragen.

Tab. 1 Vergleich der Abweichungen von Temperaturmessungen in der Stevenson-Hütte und Young-Hütte

Monat	Tagesmittelabw.	Mittel neg. Tagesabw.	Mittel pos. Tagesabw.
Jan 99	0,031	- 0,39	0,35
Feb 99	0,002	- 0,45	0,32
Mär 99	0,005	- 0,47	0,49
Apr 99	0,003	- 0,62	0,59
Mai 99	- 0,075	- 1,02	0,79
Jun 99	- 0,070	- 0,86	0,77
Jul 99	- 0,076	- 0,97	0,84
Aug 99	- 0,060	- 0,84	0,83
Sep 99	- 0,017	- 0,73	1,34
Okt 99	- 0,008	- 0,59	0,51
Nov 99	0,016	- 0,33	0,35
Dez 99	0,036	- 0,29	0,32
Jan 00	0,027	- 0,29	0,28
Feb 00	0,020	- 0,38	0,32
Mär 00	- 0,002	- 0,50	0,35
Apr 00	- 0,020	- 0,62	0,69
Mai 00	- 0,111	- 0,96	0,78
Jun 00	- 0,089	- 0,98	0,67
Jul 00	- 0,061	- 0,66	0,48
Aug 00	- 0,090	- 0,86	0,80
Sep 00	- 0,026	- 0,63	0,75
Okt 00	- 0,015	- 0,58	0,57
Nov 00	0,017	- 0,48	0,44
Dez 00	0,019	- 0,33	0,34

3. Trägheitsverhalten von ventilierten Young-Hütten

Für die ersten Feldexperimente, die dem Vergleich mit der akustischen Tomographie dienten, wurden im Eigenbau entwickelte Thermoelemente zunächst in Strahlungsschutzrohren (Schiebein und Arnold, 1996) und später in Young-Hütten eingebaut (Raabe u. a., 2000). Die aktive Lötstelle bestand aus 5 cm langen Drähten aus Konstantan und Eisen (0,2 mm Durchmesser). Die weitere Leitungsführung erfolgte in einem Aluminiumrohr 10x1 mm, welches in den Schaft der Young-Hütte eingepasst wurde. Der Wärmeübergang von der Luft an die aktive Lötstelle wurde durch einen Ventilator am oberen Ende der Young-Hütte verbessert.

Die Trägheit des Systems wird hauptsächlich vom Wärmeübergang von der Luft an den Sensor und von der Wärmeleitung über die Halterung an den Sensor bestimmt. Weiterhin spielt die Masse der Strahlungsschutzlamellen eine Rolle, durch die die Umgebungsluft strömt. Die Auflösung der Temperaturänderungen und die Trägheit des Systems waren bereits zufriedenstellend, jedoch konnten die Strahlungsfehler immer noch festgestellt werden.

Für die exakte Ermittlung der Trägheit des gesamten Systems wurden spezielle Versuche durchgeführt. Die Hütte mit dem Temperatursensor wurde in einem Klimaschrank auf eine Temperatur von etwa 7 °C abgekühlt und der Temperaturverlauf durch Aufzeichnung aller Sekunden verfolgt, nachdem die Hütte mit dem Sensor im Laborraum ohne Ventilation der Umgebung oder in einen Luftstrom mit 2,2 m/s den neuen Bedingungen ausgesetzt wurde. Aus den Temperaturverlaufskurven konnten die Trägheitsbeiwerte für die verschiedenen Phasen berechnet werden. In Abb. 1 ist der Temperaturverlauf und eine berechnete Trendlinie aus diesen Werten dargestellt.

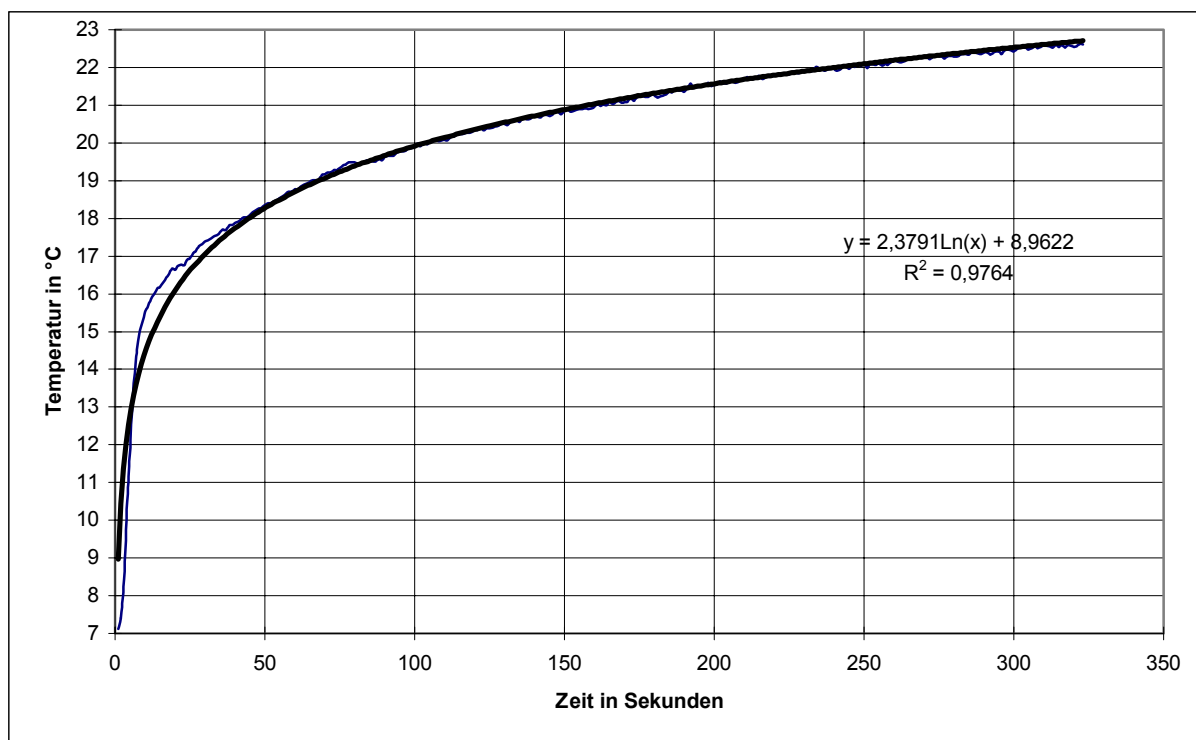


Abb. 1: Temperaturverlauf und Trendlinie (fett) bei 2,2 m/s Anströmgeschwindigkeit.

Der Verlauf der Trägheitskonstanten ist in Abb. 2 dargestellt. Die „Trägheitskonstante“ sollte für eine Messeinrichtung eine Konstante sein, in diesem Falle ist sie es nicht, da offensichtlich in der ersten Phase der Angleichung die geringe Wärmekapazität der exponierten Messdrähte und in der zweiten Phase die hohen Wärmekapazitäten der massiven Halterung und der Lamellen wirken. Die gesamte Hütte braucht einige Minuten, bis sie auch bei 2,2 m/s Anströmgeschwindigkeit die Umgebungstemperatur erreicht.

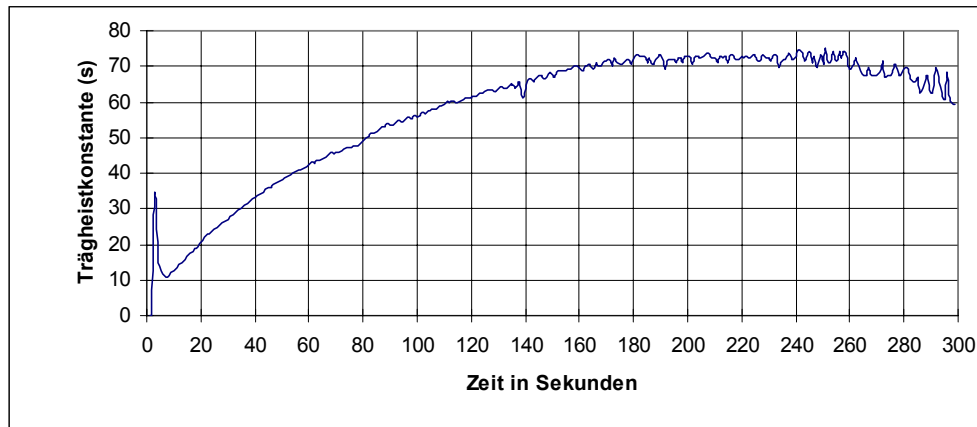


Abb. 2: Zeitlicher Verlauf der Trägheitskonstante.

4. Temperaturfehlanzeige durch die Ventilatorenwärme bei ventilierten Young-Hütten

Für den Vergleich der Messergebnisse aus der Schallausbreitung und der mit Thermoelementen ermittelten Temperaturen traten auch nachts Differenzen auf, die nicht sofort erklärt werden konnten. Eine Hypothese läuft darauf hinaus, dass die ventilierten Young-Hütten einer Eigenerwärmung durch den Ventilationslüfter unterliegen. Um diesen Sachverhalt nachzuprüfen, wurden Vergleichsversuche angestellt.

Der Temperaturverlauf von zwei in gleicher Raumhöhe in 70 cm Entfernung aufgestellter Hütten ist in Abb. 3 dargestellt. Zum Zeitpunkt 1 wurde der Ventilator der Hütte 1 in Betrieb genommen, beim Zeitpunkt 2 jener der Hütte 2. In beiden Fällen steigen die angezeigten Temperaturen um jeweils 0,3 K sofort und dann mit einer Verzögerung von einer halben Stunde um weitere 0,1 K an.

Zum Zeitpunkt 3 wurde der Ventilator der Hütte 1 wieder ausgeschaltet. Die Temperatur sinkt wieder auf den ursprünglichen Wert ab, wogegen die Temperatur an der Hütte 2, wo der Ventilator noch läuft, auf dem höheren Wert verharrt. Es dauert aber etwa 3 Stunden, bis der Ausgangswert bei der Hütte 1 erreicht worden ist (Zeitpunkt 4).

Der in Abb. 3 dargestellte Temperaturverlauf hat sich bei völliger Luftruhe im Laborraum ergeben. Interessant wäre nun, wie groß der Einfluss der Ventilationswärme bei geringen Windgeschwindigkeiten ist. Zur Ermittlung dieses Zusammenhanges wurden 2 Hütten in den Luftstrahl eines Windkanals gebracht und die Temperaturbedingungen ebenfalls über einige Zeit gemessen. Es ergeben sich die in Abb. 4 dargestellten Messwerte.

Zum Zeitpunkt 1 wurde der Ventilator von der Hütte 1 angeschaltet (Hütte 2 wurde während des gesamten Versuches unventiliert betrieben), ab 2 erfolgte die Umströmung der ventilierten Hütten im Windkanal mit 2,5 m/s und ab 3 mit 1,1 m/s. Zum Zeitpunkt 4 wurde die Ventilation der Hütte 1 ausgeschaltet. Der Temperatursprung der Hütte 2 zum Zeitpunkt 2 ergibt sich durch die Lufterwärmung bei der Inbetriebnahme des Windkanals.

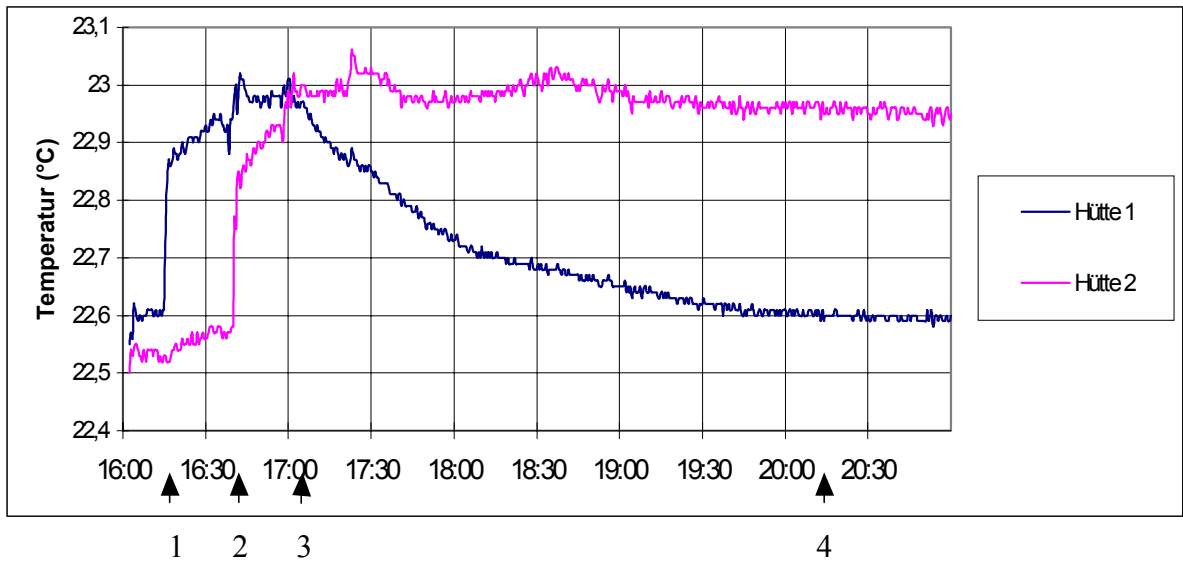


Abb. 3 : Effekte durch Ventilatorwärme.

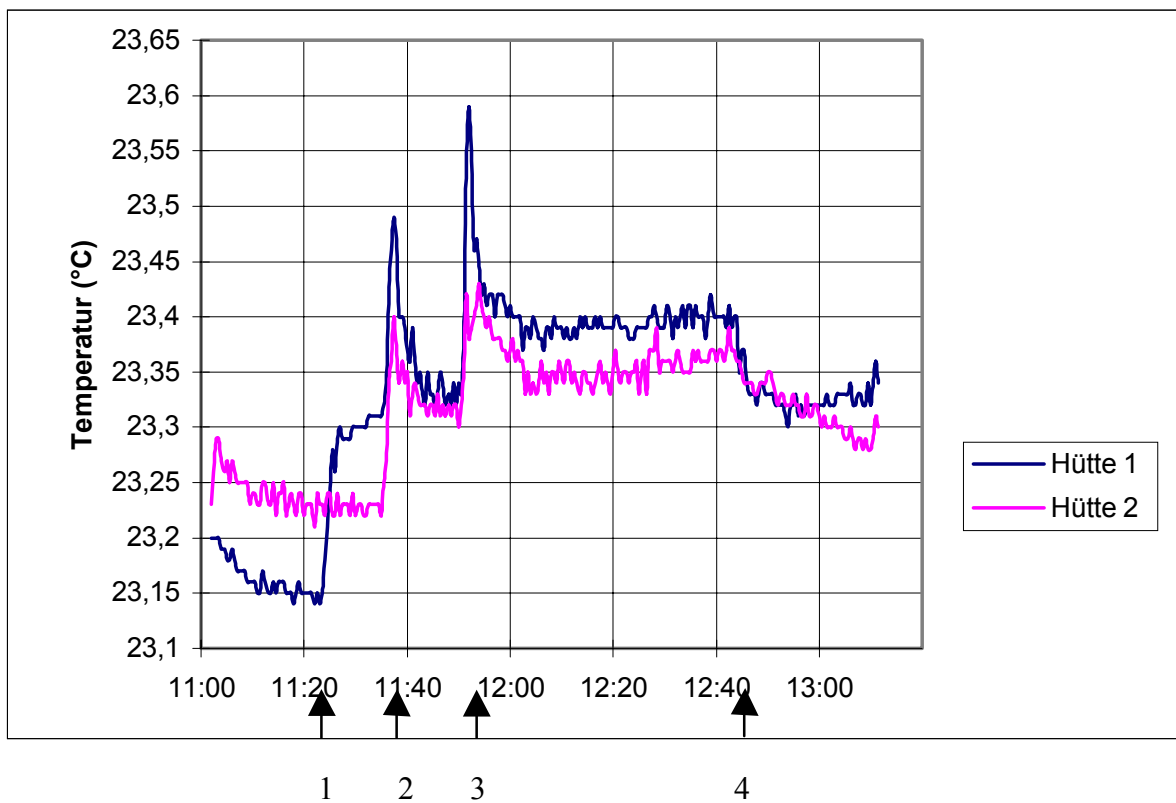


Abb. 4 : Vergleich der Hütten im Windkanal, 2 Geschwindigkeiten der Ventilation.

Aus dieser Aufzeichnung ist zu entnehmen, dass bei Ventilation mit 2,5 m/s die Unterschiede in der Anzeige verschwinden und bei 1,1 m/s gerade noch festgestellt werden können. Insgesamt muss darauf hingewiesen werden, dass diese Effekte sich im Bereich von einigen Zehntelgraden abspielen und Störungen des Temperaturverlaufs auch durch das Betreten des Laborraumes entstanden sind.

5. Bau von ventilierten Hütten

Nach Feststellung der genannten Unzulänglichkeiten bei der Temperaturmessung für die bestehende Aufgabe zum Vergleich mit den Messergebnissen der akustischen Tomographie haben sich die Autoren entschlossen, handelsübliche Sensoren mit einer ausreichenden Ventilation einzusetzen. Bei dem Aspirationspsychrometer von Assmann sind bereits 2,5 bis 3 m/s am Thermometer üblich. Diese Ventilationsgeschwindigkeit sollte auch an den HUMITTER-Fühlern der Firma VAISALA herrschen. Dafür wurde ein HUMITTER 50Y in ein poliertes Edelstahlrohr mit 40 mm Durchmesser so eingebaut, dass die Sonnenstrahlung nicht

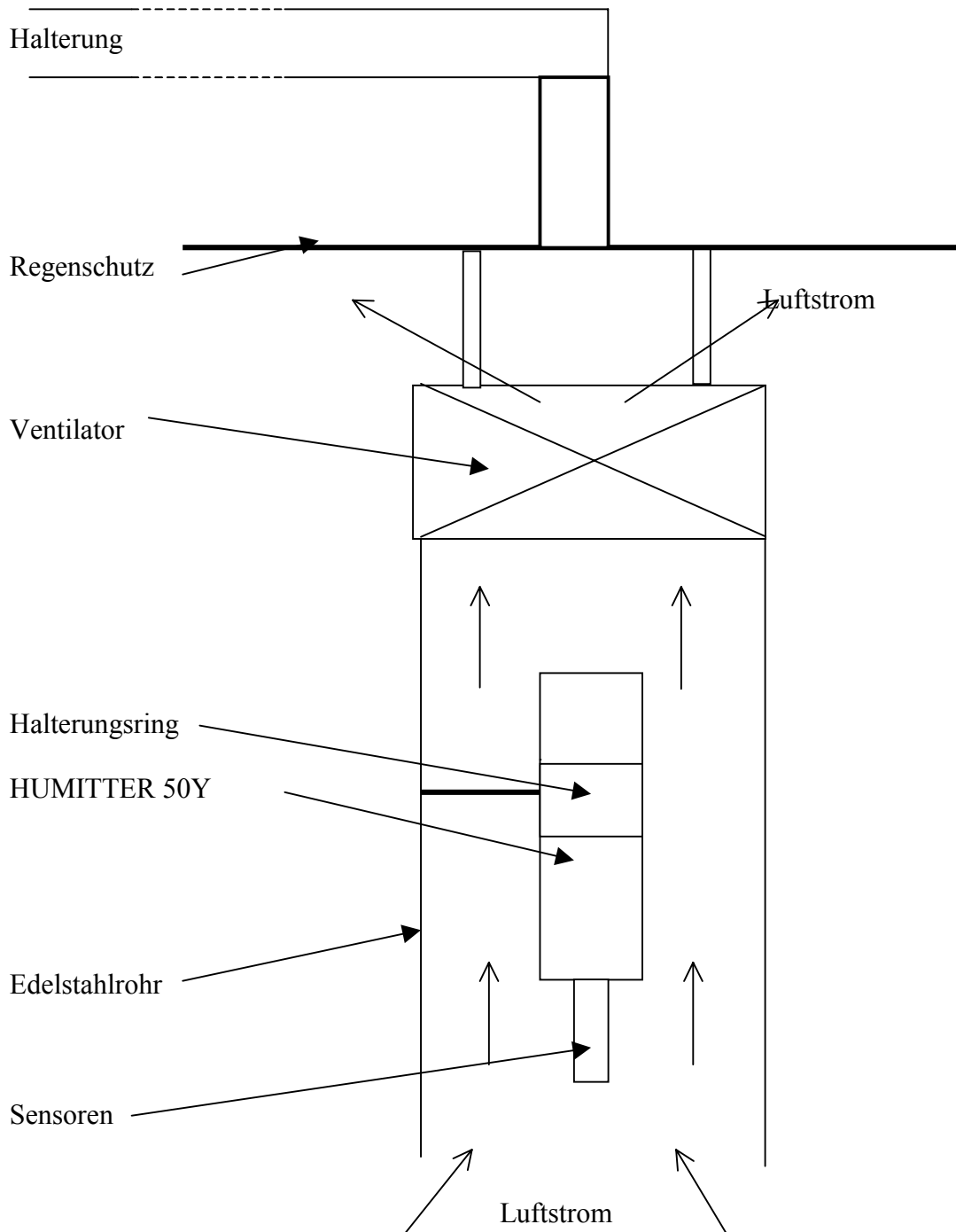


Abb. 5 : Schematischer Aufbau eines HUMITTER-Fühlers mit Ventilation.

direkt auf den Sensor treffen kann (s. Abb. 5). Am oberen Ende der Halterung ist ein Ventilator angebracht, der auch mit einer Batterie angetrieben werden kann und eine Strömungsgeschwindigkeit um den Messfühler von 3 m/s garantiert. Die Luft wird an die Umgebung befördert und die Eigenwärme nicht mehr mit dem Messfühler in Berührung gebracht.

Abb. 6 zeigt einen ventilierten HUMITTER-Fühler als Ausleger zur Erfassung der Temperatur und relativen Feuchte in 2 m Höhe über dem Boden von außen, Abb. 7 diesen Fühler von unten gesehen, wobei die Sensoren ohne Schutzkappe zu erkennen sind.



Abb. 6 : Ventilierter HUMITTER-Fühler



Abb. 7 : Ventilierter HUMITTER-Fühler - von unten gesehen

6. Untersuchungen an neuen ventilierten Hütten

6.1 Trägheit der Sensoren

Die Überprüfung der Trägheitskonstanten der Sensoren HUMITTER 50Y und die eines Thermoelementes erfolgte wie folgt. Alle Messfühler wurden in einen Klimaschrank auf Ausgangswerte von 10°C und 95 % Feuchte gebracht und dann im Laborraum bei einer Ventilation mit 3 m/s den neuen Bedingungen ausgesetzt. In den ersten Sekunden sind die Messwerte nicht zu verwenden, da durch die Umbauten keine regulären Bedingungen herrschten. In Abb. 8 wurde der zeitliche Verlauf der berechneten Trägheitswerte dargestellt. Die Trägheitskonstante des HUMITTER 50Y stellt sich für die Feuchte bei 5 Sekunden und für die Temperatur bei 7 Sekunden ein. Der Thermoelementfühler zeigt wieder eine ansteigende Tendenz auf, ist

zu Beginn der Exposition im Luftstrom bei 3 Sekunden, um nach 30 Sekunden auf 12 Sekunden anzusteigen.

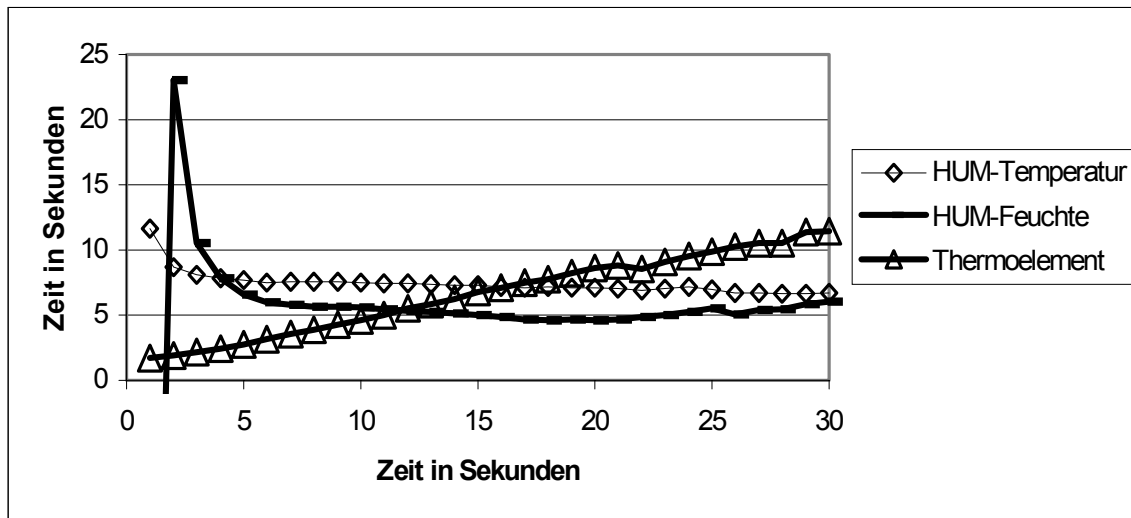


Abb. 8 : Trägheitskonstanten von HUMITTER 50Y – Fühlern und eines Thermoelementes.

6.2 Messwertvergleiche

Für die Ermittlung der Temperatur und Feuchte auf dem Versuchsgelände kommt es neben der absoluten Genauigkeit auch auf die relative an, wenn zwischen den einzelnen Messpunkten Unterschiede festgestellt werden sollen. Für die Bestimmung der absoluten Abweichungen von einem Normal-Pt100-Thermometer wurden die Geräte in einen Klimaschrank über einige Stunden bei Temperaturen von 10 bis 30 °C verglichen. Die ermittelten Korrekturen, die sich insbesondere durch die 100 m langen Übertragungsleitungen ergeben haben, sind im Datenloggerprogramm berücksichtigt worden. Danach erfolgte ein weiterer Vergleich, dessen Ergebnisse in Tabelle 2 wiedergegeben sind:

Tab. 2 Vergleich von HUMITTER 50Y-Fühlern mit Normalen

	Normal	Fühler1	Fühler 2	Fühler 3	Fühler 4
Temperatur °C	10,35	10,34	10,39	10,35	10,36
	30,10	30,09	30,08	30,08	30,05
rel. Feuchte %	38,7	39,3	38,6	38,3	39,0

Neben diesem Laborvergleich erfolgte auch eine Erprobung von 4 ventilerten Hütten über einen Zeitraum von 9 Tagen auf dem Messfeld des Instituts für Meteorologie. Dabei wurden wie für den Einsatz zum Vergleich mit der akustischen Tomographie ein Messwert pro Sekunde erfasst und Mittelwerte aller 30 Sekunden aufgezeichnet. Die aufgezeichneten Messwerte wurden untereinander verglichen. Die Mittelwerte aus den 5041 Vergleichen sind der Tabelle 3 zu entnehmen:

Tab. 3 Vergleich von ventilierten Hütten untereinander

		Hütte 1	Hütte 2	Hütte 3	Hütte 4
Temperatur	°C	18,14	18,18	18,12	18,16
rel. Feuchte	%	69,5	68,4	68,4	69,1

Aus den Vergleichen im Labor und unter Feldbedingungen, wo jeweils die vorgesehene Leitungen für die Übertragung der Messwerte zum Datenlogger in 100 m Entfernung mit eingesetzt wurden, können die Messunsicherheiten mit 0,05 K für die Temperatur und 1 % für die relative Feuchte angegeben werden.

7. Zusammenfassung

Die Verwendung von trägheitsarmen Sensoren mit geringen Wärmeübergängen an den Halterungen und die Sicherstellung einer ausreichenden zielgerichteten Ventilation lässt die messtechnische Erfassung der Lufttemperatur und der relativen Feuchte auf einem Versuchsfeld mit speziell kalibrierten Sensoren zu, deren Fehler unter einem Zehntelgrad bei einem Temperaturvergleich und 1 % bei der relativen Feuchte betragen. Die Trägheit ist geringer als bei den bisher eingesetzten Messfühlern. Der Strahlungsfehler ist durch die unzureichende Oberflächenbehandlung leider immer noch vorhanden, ein doppelter Strahlungsschutz wie beim Aspirationspsychrometer könnte auch diesen Fehler noch vermindern.

Literatur

Raabe, A., Arnold, K. & Ziemann, A. 2000: Akustische Tomographie auf dem GM Falkenberg. *Abschlußbericht Inst. f. Meteorologie der Universität Leipzig, März 2000*

Schienbein, S. & Arnold, K. 1996: Konzeption und Aufbau einer Meßeinrichtung für die Erfassung von Temperaturen und Strömungsgeschwindigkeiten bis 12 m Höhe. *Wiss. Mitt. Inst. f. Meteorol. Univ. Leipzig u. Institut f. Troposphärenforschung Leipzig, 4*, 124-130

Adresse der Autoren:

Sigurd Schienbein, Klaus Arnold
Universität Leipzig, Institut für Meteorologie
Stephanstr. 3
04103 Leipzig

Zur Schätzung von Häufigkeitstrends von extremen Wetter- und Klimaereignissen

M. Mudelsee, M. Börngen und G. Tetzlaff

Zusammenfassung

Die Vorteile der Kernschätzung gegenüber dem Abzählen von Ereignissen in Zeitintervallen werden dargestellt. Für das beiden Methoden gemeinsame Glättungsproblem gestattet die Kreuzvalidierung eine Lösung. Für die Hochwasserereignisse der Oder im Zeitraum 1350 bis 1850 wird eine Abnahme der Häufigkeit nach ca. 1675 gefunden; weitergehende Aussagen bedingen eine Homogenisierung der Daten. Die dargestellte Methodik wird gegenwärtig in das Computerprogramm XTREND implementiert.

Summary: On the estimation of trends in the frequency of extreme weather and climate events

The advantages of kernel estimation over counting of events within time intervals are shown. Cross validation offers a solution for the smoothing problem which is common to both methods. As regards flooding events of the river Oder in 1350 to 1850, a decrease in the frequency after about 1675 is found. More detailed results demand homogenized data. The method is currently being implemented into the computer program XTREND.

1. Einleitung

„Natural Hazards“ (Bryant 1991) ist dabei, ein Schlagwort zu werden vergleichbar mit „Global Change“ (Houghton *et al.* 1996). Solche extremen Wetter- oder Klimaereignisse treten mannigfaltig auf: Wirbelstürme, Starkregen, Hochwasser, Dürreperioden usw. Da sie zum Teil ein enormes Zerstörungspotential aufweisen, ist die Frage nach ihrer natürlichen Variabilität und einem möglichen anthropogenen Einfluß von gesellschaftlicher Bedeutung.

Eine mögliche Analysemethode wird durch die *Extreme Value Theory* (Embrechts *et al.* 1997) bereitgestellt: Schätzung der Wahrscheinlichkeitsverteilung von zum Beispiel Dollar-Schäden durch Hurrikane im Südosten der USA. Besondere Bedeutung dabei haben die Schwänze der Verteilung. Als ein Resultat wird die Wahrscheinlichkeit bestimmt, daß in einem Zeitbereich ein Mindestschaden auftreten wird. Grundannahme dieser Methode ist Stationarität, d. h. eine zeitlich konstante Verteilung. Für genügend kurze Beobachtungszeiträume (jedoch länger noch als der Jahreszyklus), und eventuell nach Durchführung geeigneter Transformationen (im Falle der Hurrikan-Dollar-Schäden, Zeitraum 1925–95, die Division durch den zeitabhängigen ökonomischen Reichtum im Gebiet (Pielke und Landsea 1998)), mag die Stationaritätsannahme hinreichend begründet sein.

Für längere Zeiträume, wie von den Anfängen der Industrialisierung bis heute, ist

die Annahme einer konstanten natürlichen Variabilität extremer Wetter- und Klimaereignisse unrealistisch. Eine kausale Verknüpfung mit anthropogenem Einfluß bedingt die Möglichkeit einer Änderung der Variabilität, d. h. eines **Trends in der Häufigkeit**. Zudem werden extreme Wetter- und Klimaverhältnisse zuweilen nur als Ja/Nein-Information archiviert (z. B. Frostringe in Bäumen), die Indikator-Qualität der gemessenen Klimavariablen (z. B. Sauerstoffisotopenverhältnisse in Eisbohrkernen als Temperaturanzeiger) ist eingeschränkt, oder die Daten (wie die hier untersuchten) bestehen aus historischen Aufzeichnungen und zeigen entsprechende subjektive Einflüsse — eine Schätzung der Wahrscheinlichkeitsdichte ist dann nicht möglich oder stark fehleranfällig. Deshalb ist die Analyse von Trends in der Häufigkeit unerlässlich zur Beurteilung der natürlichen Variabilität extremer Wetter- und Klimaereignisse.

Die einfachste Methode, Häufigkeitstrends abzuschätzen, ist bloßes Abzählen der Extremereignisse innerhalb von Zeitintervallen, entsprechend der Histogrammbildung bei der Wahrscheinlichkeitsdichteschätzung. Die Nachteile dieser Methode (Positionierung der Intervallgrenzen, Informationsverlust, „holpriges“ Aussehen) werden in Abschnitt 2 dargestellt. Diese Nachteile werden, analog zur Dichteschätzung, durch die Kernschätzung (Abschnitt 3) aufgehoben. Es scheint so (Literaturrecherche im *Science Citation Index* 1999), daß die Kernschätzung, obwohl Standardwerkzeug in der Statistik seit über 20 Jahren, noch keinen Eingang in meteorologische/klimatologische Anwendungen gefunden hat — mit Ausnahme von Solow (1991), der die Auftrittshäufigkeit von Vulkaneruptionen untersucht.

Von entscheidender Bedeutung ist, auch bei der Kernschätzung, das Glättungsproblem (Länge der Zeitintervalle), das hier mit Hilfe der Kreuzvalidierung (Abschnitt 4) gelöst wird. Illustriert werden die Schätzverfahren durch die Zeitreihe von Hochwasserereignissen für die Oder im Zeitraum 1350 bis 1850. Es zeigt sich (Abschnitt 5), daß ab ca. 1675 die Häufigkeit von Hochwasserereignissen der Oder abnahm.

2. Histogrammschätzung

Abbildung 1 zeigt die Daten (unter Berücksichtigung der Kalenderumstellung) von Hochwasserereignissen der Oder im Zeitraum 1350 bis 1850 nach den Weikinnischen Quellentexten (Weikinn 1958–1963, 2000, o. J.).

Abbildung 2 zeigt als Histogrammschätzung die Hochwasseranzahl in 50-Jahres-Intervallen, beginnend mit dem Jahr 1350. Abbildung 3 zeigt die Hochwasseranzahl ebenfalls in 50-Jahres-Intervallen, jedoch um 25 Jahre versetzt, beginnend 1375. Während man Abb. 2 vielleicht so beschreiben würde: „Anstieg von 1350 bis 1550, danach ungefähr konstant bzw. leicht ansteigend, mit Ausnahme der deutlichen Erhöhung im Bereich 1650–1700“, könnte man Abb. 3 deuten als: „Anstieg von 1375 bis 1650, danach leichter Abfall“. Diese Zweideutigkeit, hervorgerufen durch die Positionierung der Intervallgrenzen, erschwert offensichtlich eine vernünftige Interpretation des Häufigkeitstrends.

Weiterhin ist festzustellen, daß die Reduktion der Datenanzahl ($N = 169$) auf die Anzahl der Zeitpunkte, an der eine Histogrammschätzung vorliegt (d. h. die Klassenanzahl $n_k = 10$ bzw. 9) einen Informationsverlust darstellt (Abb. 2, 3).

Darüberhinaus führt die Histogrammschätzung aufgrund der Ja/Nein-Möglichkeit der Intervallzugehörigkeit generell zu einem „holprigen“ Aussehen der geschätzten Häufigkeit, wie hier besonders in Abbildung 4 deutlich wird. Obwohl das als von „kosmetischer

Natur“ (Diggle 1985) empfunden werden kann, ist es dennoch ein „unnatürlicher“ Aspekt.

Von entscheidender Bedeutung ist jedoch die Wahl der Intervallbreite, das Glättungsproblem. Abbildungen 2 bis 5 verwenden die gleichen Daten, jedoch unterschiedliche Intervallbreiten. Eine große Breite (Abb. 5) bedeutet mehr Ereignisse pro Intervall und damit einen niedrigeren statistischen Fehler der geschätzten Häufigkeit. Gleichzeitig können kurzzeitigere Datenmerkmale nicht mehr aufgelöst werden (erhöhter systematischer Fehler). Umgekehrt im Falle einer kleinen Intervallbreite (Abb. 4): erhöhter statistischer Fehler bei niedrigerem systematischem Fehler.

Eine Möglichkeit zur Lösung des Glättungsproblems, des Ausbalancierens des statistischen gegen den systematischen Fehler, wird in Abschnitt 4 dargestellt.

3. Kernschätzung

Es seien T_1, T_2, \dots, T_N die geordneten Beobachtungen von Extremereignissen im Zeitintervall $[T_S; T_E]$. Die Daten werden als Realisierungen eines nichthomogenen Poisson-Prozesses (Cox und Lewis 1966) betrachtet:

1. die Wahrscheinlichkeit, daß in einem infinitesimalen Zeitintervall $[t; t + \delta]_{\delta \rightarrow 0} \in [T_S; T_E]$ ein Extremereignis auftritt, ist gleich $\delta \cdot \lambda(t)$;
2. die Wahrscheinlichkeit gleichzeitiger Ereignisse ist vernachlässigbar;
3. die Wahrscheinlichkeiten beobachteter Ereignisse in sich nicht überschneidenden Intervallen aus $[T_S; T_E]$ sind unabhängig voneinander.

$\lambda(t)$ wird Auftrittsrates oder in der Statistik üblicherweise Intensität genannt. Die Aufgabe ist also, die Intensität aus den Daten der Extremereignisse zu schätzen.

Im Fall der Hochwasserdaten (Abb. 1) sollte Eigenschaft 2 ausreichend erfüllt sein: bei der Auswertung der Weikinnschen Quellensammlung wurde darauf geachtet, daß mehrere Meldungen zu einem einzelnen Hochwasserereignis, etwa an aufeinander folgenden Tagen, auch nur einen Dateneintrag erhalten. Ebenfalls Eigenschaft 3 sollte gut erfüllt sein. Man müsste schon abenteuerliche Gegenkonstruktionen errichten, etwa: „Das Oder-Jahrhunderthochwasser 1736 [vgl. dazu Miltzer *et al.* (1999)] dezimierte die ansässige Bevölkerung so stark, daß in den Folgejahren keine Hochwasserbeobachtungen mehr durchgeführt werden konnten...“

Die Kernschätzung (Rosenblatt 1956, Diggle 1985) schätzt die Intensität $\lambda(t)$ als

$$\hat{\lambda}_h(t) = \sum_{i=1}^N K_h(t - T_i), \quad t \in [T_S; T_E], \quad K_h(\cdot) = h^{-1}K(\cdot/h) \quad (1)$$

mit dem Glättungsparameter oder Bandbreite h . Die Wahl der genauen Form der Kernfunktion K ist von sekundärer Bedeutung, solange K symmetrisch und ≥ 0 ist und $\int K = 1$ ist. Wir wählen eine Gauß-Funktion, weil sich $\hat{\lambda}_h$ dann recheneffektiv im Fourierraum (mit *Fast Fourier Transform*, FFT) bestimmen läßt (Silverman 1982). Das sanfte Auslaufen der Gauß-Funktion nach $\pm\infty$ beseitigt auch den dem Histogramm (K : Rechteckfunktion) inhärenten Nachteil des „holprigen“ Aussehens von $\hat{\lambda}$ (Abschnitt 2).

In Gleichung (1) kann $\hat{\lambda}_h$ an beliebig fein eingestellten t berechnet werden, was die Willkürlichkeit der Positionierung der Klassengrenzen sowie den Informationsverlust der

Histogrammschätzung (Abschnitt 2) aufhebt. Ein Gitter von 1024 Punkten (Potenz von 2 wegen der FFT) ergibt eine gute visuelle Darstellung bei geringer Rechenzeit (wenige Sekunden bei dem hier untersuchten Beispiel ($N = 169$) und einem 433-MHz-Celeron-PC).

Zur Wahl des Glättungsparameters h siehe den folgenden Abschnitt. Es sei noch auf den Vorteil der Kernschätzung als nichtparametrische Schätzmethode verwiesen: man kommt ohne strukturelle Annahmen aus, wie sie zum Beispiel das Modell eines linearen Anstiegs der Häufigkeit mit der Zeit tut. Das bedeutet, daß die Kernschätzung von Häufigkeitstrends ein flexibles Verfahren ist, einsetzbar auf unterschiedlichen meteorologischen und klimatologischen Anwendungsfeldern (Abschnitt 1).

4. Glättungsproblem

Die Wahl des Glättungsparameters h beeinflusst (Gleichung 1, siehe auch Abschnitt 2 sowie Abb. 7 und 8) entscheidend die Form des geschätzten Intensitätsverlaufes, $\hat{\lambda}_h(t)$. Es wird eine Bandbreite gesucht, für die, einfach gesprochen, $\hat{\lambda}_h(t)$ „nahe bei“ $\lambda(t)$, dem wahren Intensitätsverlauf liegt bzw. ein ausgewogenes Verhältnis statistischer/systematischer Fehler vorliegt. Es gibt unterschiedliche Maße für das „nahe bei“; wir folgen hier Brooks und Marron (1991), die den *integrated square error* (*ISE*) heranziehen:

$$\begin{aligned} ISE_\lambda(h) &= \int_{T_S}^{T_E} \left[\hat{\lambda}_h(t) - \lambda(t) \right]^2 dt \\ &= \int_{T_S}^{T_E} \hat{\lambda}_h(t)^2 dt - 2 \int_{T_S}^{T_E} \hat{\lambda}_h(t) \lambda(t) dt + \int_{T_S}^{T_E} \lambda(t)^2 dt. \end{aligned} \quad (2)$$

Brooks und Marron geben weiter die Kreuzvalidierungsfunktion an als

$$CV_\lambda(h) = \int_{T_S}^{T_E} \hat{\lambda}_h(t)^2 dt - 2 \sum_{i=1}^N \hat{\lambda}_{h,i}(T_i), \quad (3)$$

worin

$$\hat{\lambda}_{h,i}(t) = \sum_{\substack{j=1 \\ j \neq i}}^N K_h(t - T_j) \quad (4)$$

der *leave-one-out*-Kernschätzer ist. Analog zu der Schätzung des Erwartungswertes $\mu = \int x f(x) dx$ einer Wahrscheinlichkeitsdichte f durch den Mittelwert einer Stichprobe, $\bar{x} = 1/n \sum_{i=1}^n x_i$, läßt sich $\sum_{i=1}^N \hat{\lambda}_{h,i}(T_i)$ als Momentenschätzer von $\int_{T_S}^{T_E} \hat{\lambda}_h(t) \lambda(t) dt$ auffassen. Das Weglassen des Summanden $j = i$ in Gleichung (4) sorgt dafür, daß aus der gemeinsamen Stichprobe $\{T_1, \dots, T_N\}$ nicht sowohl die Schätzung als auch das Maß (*ISE*)

der Schätzqualität genommen werden. Da $\int_{T_S}^{T_E} \lambda(t)^2 dt$ unabhängig von h ist, liegt diejenige Bandbreite h_{CV} , die $CV_\lambda(h)$ minimiert, nahe bei der Bandbreite h_{ISE} , die $ISE_\lambda(h)$ minimiert. (Die Konvergenz jedoch, mit der h_{CV} gegen h_{ISE} strebt, geht sehr langsam mit N , vgl. Marron (1987).)

Abbildung 6 zeigt die Kreuzvalidierungsfunktion für die Wahl des Glättungsparameters im Falle der Hochwasserereignisse der Oder im Zeitraum 1350 bis 1850. Das Minimum liegt bei $h \approx 20$ Jahre. Abbildung 7 zeigt die entsprechende Kernschätzung des Intensitätsverlaufes.

Wegen der oben erwähnten langsamen Konvergenz von $h_{CV} \rightarrow h_{ISE}$ sollte bei dem Oderhochwasser-Datensatz ($N = 169$) auch mit anderen Bandbreiten experimentiert werden. Als obere Grenze für h würde vielleicht $h = 75$ Jahre akzeptabel sein, wofür $CV_\lambda(h)$ noch nicht stark gegenüber dem Minimum angewachsen ist (Abbildung 6). Abbildung 8 zeigt die entsprechende Kernschätzung des Intensitätsverlaufes.

Durch die Kreuzvalidierung wird also zumindest ein geeigneter Bereich für h vorgegeben. Die letztliche Auswahl von h hat deshalb subjektive Anteile. Für die hier untersuchten Oderhochwasser erscheint $h = 20$ Jahre vielleicht etwas zu klein, entsprechende hydrologische/meteorologische Änderungen dürften etwas langsamer ablaufen. Marron (1988) gibt weiter an, daß die Kreuzvalidierungsfunktion häufig auch lokale Minima zeigt (im Unterschied zu Abb. 6 hier) und deshalb eine *brute-force*-Berechnung von $CV_\lambda(h)$ angezeigt ist. Das Computerprogramm XTREND, mit dem die Kernschätzungen hier durchgeführt wurden, ist deshalb für ein interaktives Arbeiten ($CV_\lambda(h)$ grafisch, Auswahl von h , Plot von $\hat{\lambda}_h(t)$, Neuauswahl von h usw.) ausgerichtet.

5. Ergebnis

Ein Problem bei der Interpretation der geschätzten Intensitätsverläufe der Oder-Hochwasserereignisse (Abb. 7 und 8) ist die Inhomogenität der Daten. Es ist davon auszugehen, daß im untersuchten Zeitraum 1350–1850 die Wahrscheinlichkeit einer Hochwasser-Aufzeichnung, vorgenommen durch natur- und überlieferungsinteressierte Menschen, zugenommen hat — eventuell mit einer Unterbrechung im Dreißigjährigen Krieg 1618–48. Eine realistischere, in Zukunft vorzunehmende Datenanalyse erfordert deshalb als ersten Schritt eine entsprechende Homogenisierung. Diese kann einerseits durchgeführt werden mit Hilfe von Zählungen meteorologischer Mitteilungen wie Hellmann (1883); diese Arbeit erfaßt jedoch nur den Bereich 1484–1881. Sofern die Annahme berechtigt ist, daß die Aufzeichnungswahrscheinlichkeit sich langsamer geändert hat (Größenordnung > 100 Jahre) als das Hochwassergeschehen ($h = 20$ – 75 Jahre, siehe Abb. 7–8), könnte eine Homogenisierung erfolgen über Normierung:

$$\hat{\lambda}_{h_{CV}}(t) / \hat{\lambda}_{h \gtrsim 100 \text{ Jahre}}(t).$$

Weiterhin sollen ebenfalls unterschiedliche Stärkeklassen der Hochwasser, ermittelt aus Anzahl ertrunkener Menschen, Ausmaß der ökonomischen Schäden, Anzahl der Meldungen usw., untersucht werden. Eine Differenzierung in Sommer-/Winterhochwasser mag Zusatzinformation erbringen. Schließlich sollen die Oderhochwasser im Zeitraum 1850 bis heute (welche nicht in den Weikinnschen Quellentexten enthalten sind), einbezogen werden.

Trotz der anzunehmenden Inhomogenität der Daten (s. o.) kann man jedoch festhalten (Abb. 7, 8), daß nach ca. 1675 die Häufigkeit von Hochwasserereignissen der Oder abnahm. Eine Homogenisierung würde diesen Befund lediglich verstärken.

Um weitergehende Aussagen machen zu können, sind neben der Datenhomogenisierung und der Erweiterung des Zeitbereiches folgende methodische Ergänzungen geplant:

1. Pseudodatenerzeugung (Cowling und Hall 1996) im Bereich vor 1350, um Randeffekte zu mildern;
2. *Bootstrap*-Simulationen (Cowling *et al.* 1996) zur Ermittlung von Konfidenzbändern von $\hat{\lambda}_h(t)$. Erst damit kann beantwortet werden, wie signifikant die Häufigkeitsabnahme seit 1675 ist, oder ob kurzzeitige Änderungen im Hochwassergeschehen der Oder, angezeigt durch den Intensitätsverlauf bei kleiner Bandbreite h_{CV} (Abb. 7), „real“ sind.

6. Schlussfolgerungen

Die Kernschätzung ist besser als die Histogrammschätzung geeignet, Häufigkeitstrends von extremen Wetter- und Klimaereignissen zu schätzen, weil sie (1) kein Positionierungsproblem der Histogrammgrenzen aufweist, (2) keinen Informationsverlust bewirkt und (3) die geschätzte Auftrittsrates oder Intensität kein „unnatürliches holpriges“ Aussehen zeigt (bei Gaußscher Kernfunktion). Wie bei der Histogrammschätzung, so ist auch bei der Kernschätzung das Glättungsproblem zentral. Kreuzvalidierung gibt eine Lösungsorientierung, in der konkreten Analyse ist das Experimentieren mit unterschiedlichen Bandbreiten jedoch unverzichtbar.

Das Ergebnis, der Häufigkeitstrend von Hochwasserereignissen der Oder im Zeitraum 1350 bis 1850, trägt lediglich vorläufigen Charakter, weil (1) die Daten nicht homogenisiert sind (Aufzeichnungswahrscheinlichkeit) und (2) noch keine *Bootstrap*-Konfidenzbänder von dem geschätzten Intensitätsverlauf ermittelt wurden. Es scheint so, daß ab ca. 1675 die Hochwasserhäufigkeit für die Oder, wie aufgezeichnet durch die Weikinnschen Quellentexte, abnahm.

Die hier dargestellte Methodik (Kernschätzung, Pseudodatenerzeugung, *Bootstrap*-Simulationen) wird gegenwärtig in das Computerprogramm XTREND implementiert.

Dank

Herrn Dr. H.-J. Schönfeldt sei für die kritische Rückfrage zur konkreten FFT-Berechnung herzlich gedankt. Cand. met. Korinna König digitalisierte die verwendete Oderhochwasser-Zeitreihe aus den Weikinnschen Quellentexten. Die vorliegende Arbeit wurde finanziell unterstützt durch die Deutsche Forschungsgemeinschaft (Mu 1595/1-1, Mu 1595/1-2, Te 51/15-1).

Literatur

Brooks MM, Marron JS (1991) Asymptotic optimality of the least-squares cross-validation bandwidth for kernel estimates of intensity functions. *Stochastic Processes and their Applications* 38:157–165.

- Bryant E (1991) *Natural Hazards*. Cambridge University Press, Cambridge, 294 S.
- Cowling A, Hall P (1996) On pseudodata methods for removing boundary effects in kernel density estimation. *Journal of the Royal Statistical Society Series B* 58:551–563.
- Cowling A, Hall P, Phillips MJ (1996) Bootstrap confidence regions for the intensity of a Poisson point process. *Journal of the American Statistical Association* 91:1516–1524.
- Cox DR, Lewis PAW (1966) *The Statistical Analysis of Series of Events*. Methuen, London, 285 S.
- Diggle P (1985) A kernel method for smoothing point process data. *Applied Statistics* 34:138–147.
- Embrechts P, Klüppelberg C, Mikosch T (1997) *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, 648 S.
- Hellmann G (1883) *Repertorium der Deutschen Meteorologie*. Verlag von Wilhelm Engelmann, Leipzig, 996 S.
- Houghton JT, Meira Filho LG, Callander BA, Harris N, Kattenberg A, Maskell K (1996) *Climate Change 1995: The Science of Climate Change*. Cambridge University Press, Cambridge, 572 S.
- Marron JS (1987) What does optimal bandwidth selection mean for nonparametric regression estimation? In: Dodge Y (Hrsg.) *Statistical Data Analysis Based on the L₁-Norm and Related Methods*. North-Holland, Amsterdam, 379–392.
- Marron JS (1988) Automatic smoothing parameter selection: A survey. *Empirical Economics* 13:187–208.
- Militzer S, Börngen M, Tetzlaff G (1999) Das Oderhochwasser von 1736: Ein Analogon zum Jahrhunderthochwasser 1997. *Geoökodynamik* 20:309–322.
- Pielke RA Jr, Landsea CW (1998) Normalized hurricane damages in the United States: 1925–95. *Weather and Forecasting* 13:621–631.
- Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics* 27:832–837.
- Silverman BW (1982) Kernel density estimation using the Fast Fourier Transform. *Applied Statistics* 31:93–99.
- Solow AR (1991) An exploratory analysis of the occurrence of explosive volcanism in the Northern Hemisphere, 1851–1985. *Journal of the American Statistical Association* 86:49–54.
- Weikinn C (1958–1963) *Quellentexte zur Witterungsgeschichte Europas von der Zeitwende bis zum Jahre 1850, Hydrographie, Teile 1–4*. Akademie-Verlag, Berlin.
- Weikinn C (2000) *Quellentexte zur Witterungsgeschichte Europas von der Zeitwende bis zum Jahre 1850, Hydrographie, Teil 5*. Herausgegeben und bearbeitet von Michael Börngen und Gerd Tetzlaff. Borntraeger-Verlag, Berlin.
- Weikinn C (o. J.) *Quellentexte zur Witterungsgeschichte Europas von der Zeitwende bis zum Jahre 1850, Hydrographie, Teil 6*. Herausgegeben und bearbeitet von Michael Börngen und Gerd Tetzlaff. Borntraeger-Verlag, Berlin. [in Bearbeitung]

Adresse der Autoren:

Dr. Manfred Mudelsee, Dr. Michael Börngen, Prof. Dr. Gerd Tetzlaff
 Universität Leipzig, LIM — Institut für Meteorologie, Stephanstr. 3, 04103 Leipzig

Hochwasser-Ereignisse Oder [1350-1850]

(N = 169)

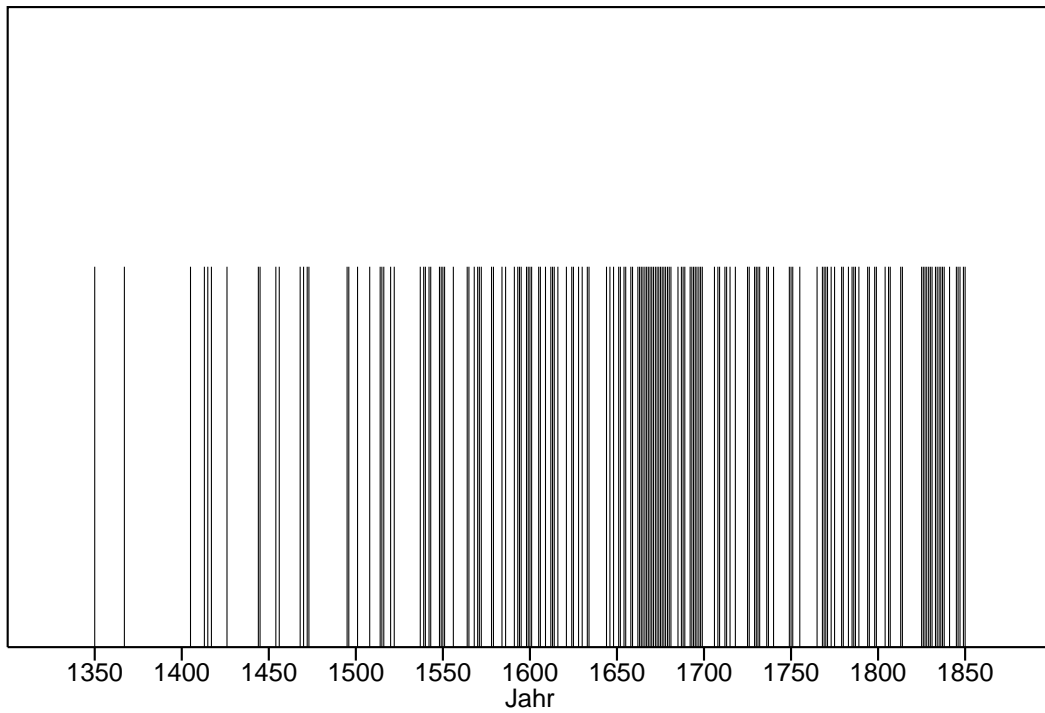


Abbildung 1: Untersuchte Daten.

Hochwasser-Ereignisse Oder [1350-1850]

(N = 169)

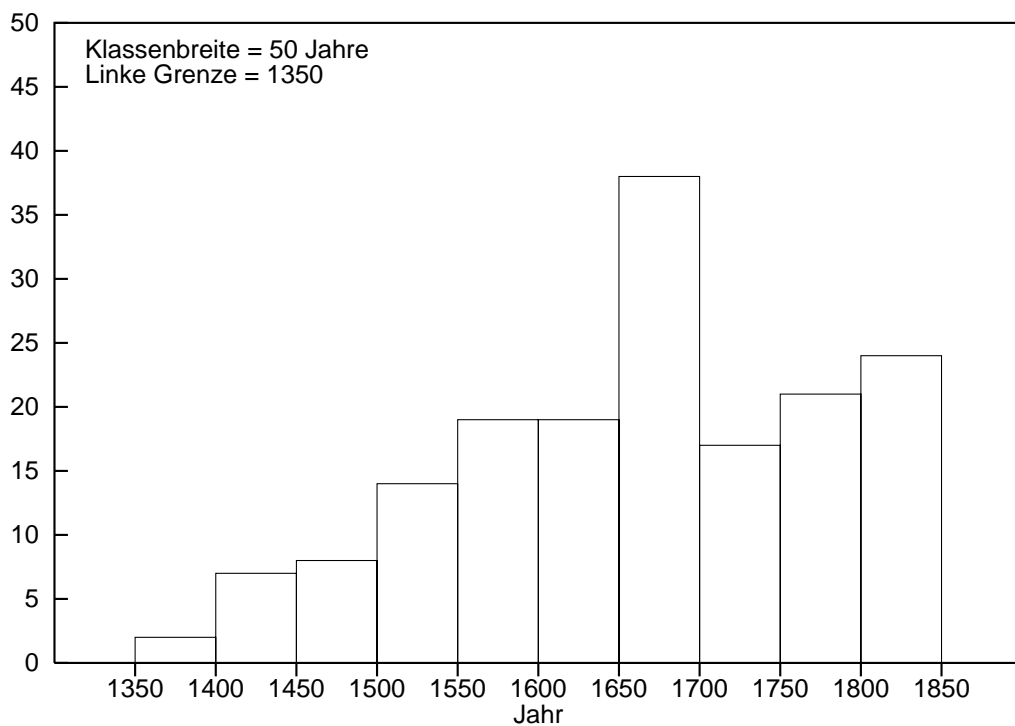


Abbildung 2. Histogrammschätzung. Die mathematischen Histogrammintervall-Grenzen: [[[...]].

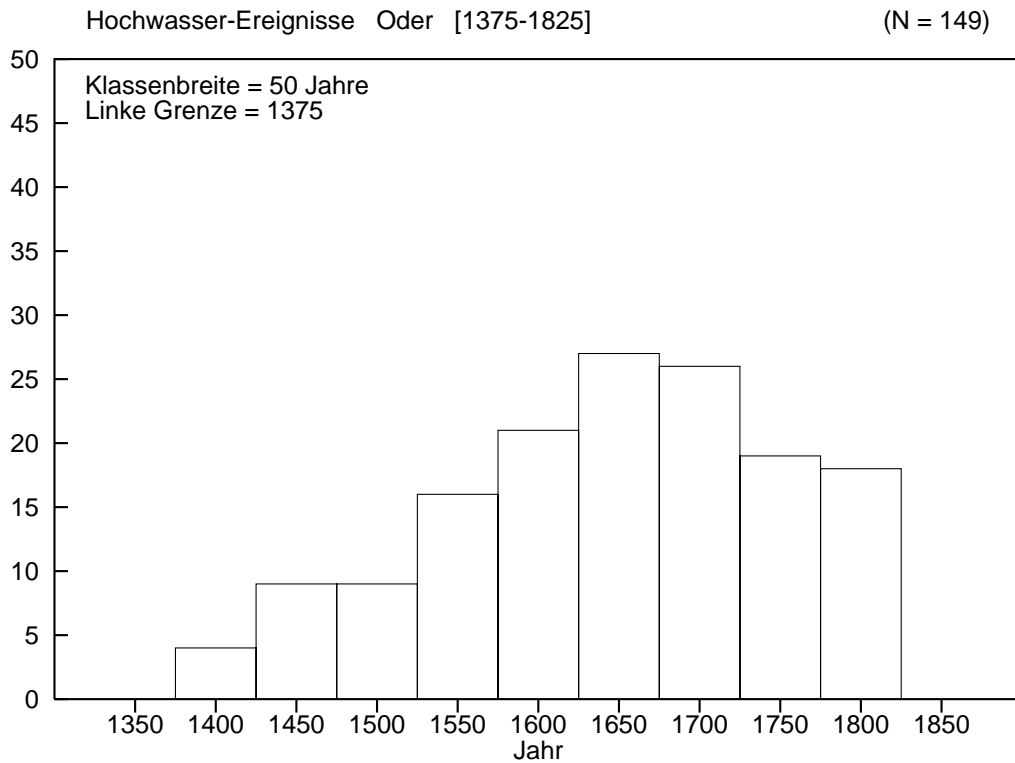


Abbildung 3. Histogrammschätzung. Die Daten in den Bereichen [1350; 1375[und]1825; 1850] sind nicht dargestellt.

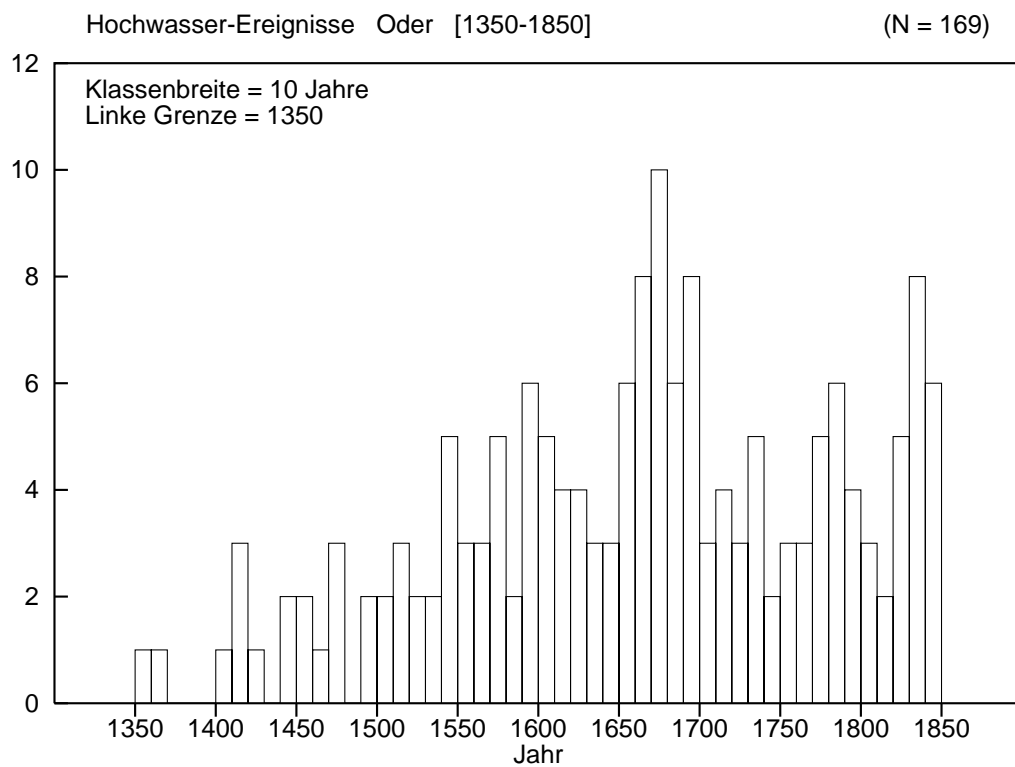


Abbildung 4: Histogrammschätzung.

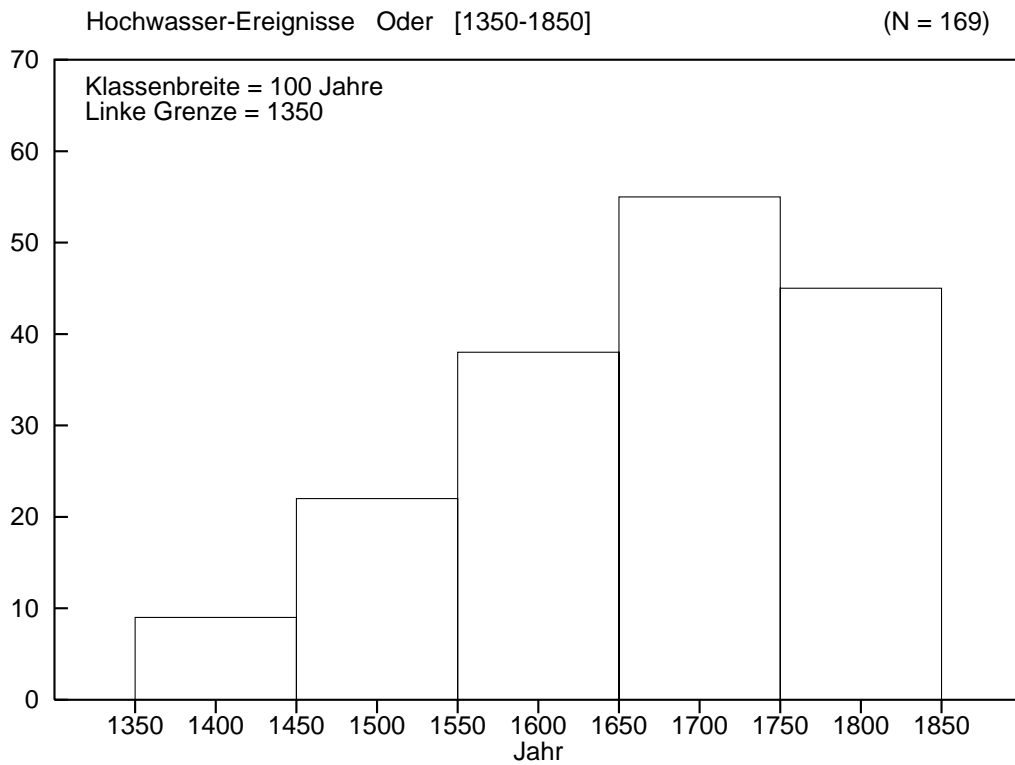


Abbildung 5: Histogrammschätzung.



Abbildung 6. Kreuzvalidierungsfunktion $CV_{\lambda}(h)$. Es wurde $T_S = T_1$ und $T_E = T_N$ angesetzt. Das Minimum liegt bei $h_{CV} = 20$ Jahre.

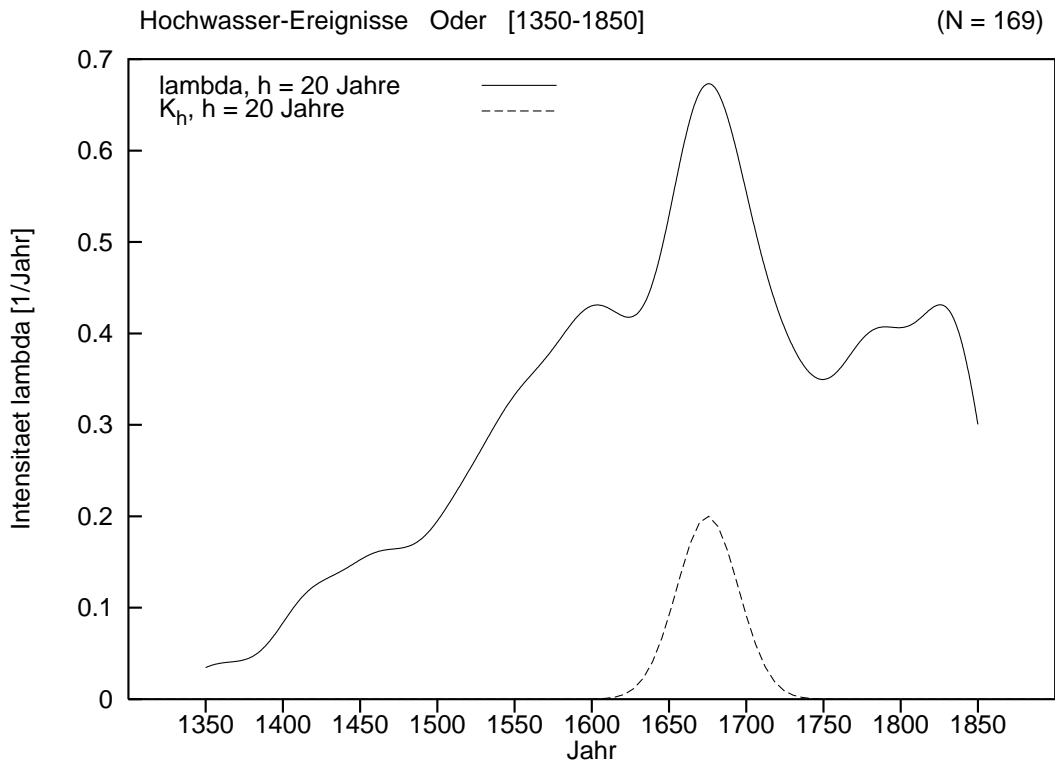


Abbildung 7. Kernschätzung des Intensitätsverlaufes. Zur Veranschaulichung ist auch die Kernfunktion K_h eingezeichnet.

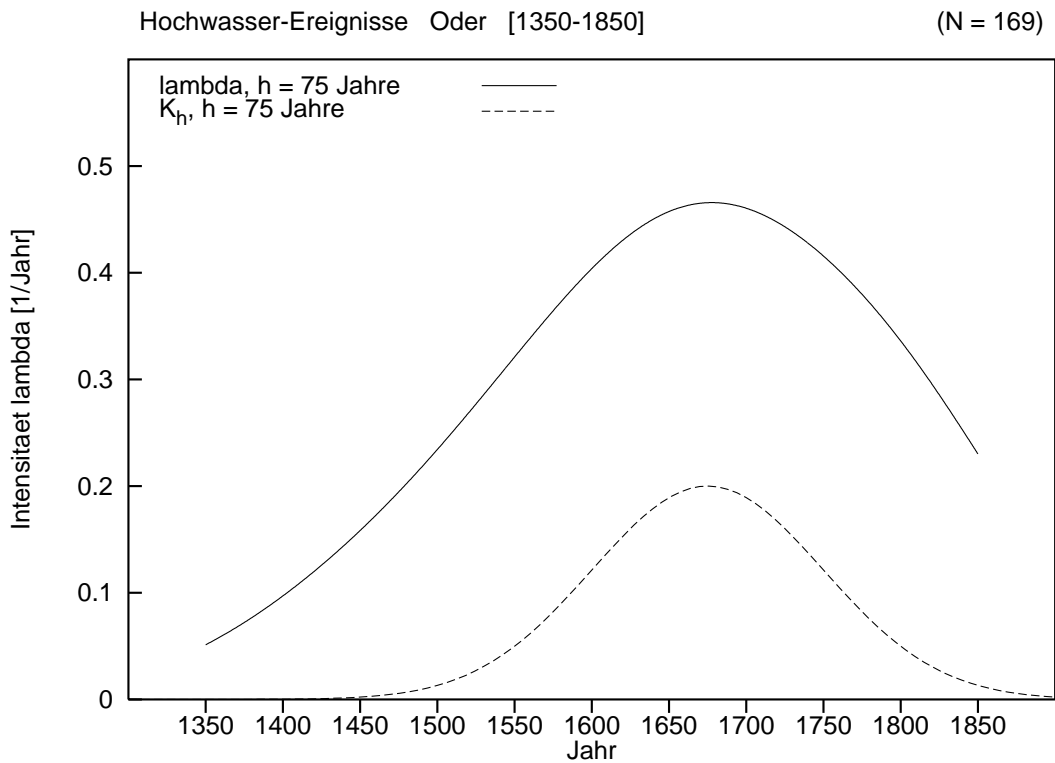


Abbildung 8. Kernschätzung des Intensitätsverlaufes. Zur Veranschaulichung ist auch die Kernfunktion K_h eingezeichnet.

Datenfenster und deren Wirkung in der spektralen Varianzanalyse geophysikalischer Meßreihen

H.-J. Schönfeldt und S. von Löwis

Zusammenfassung

Die ständig steigende Rechenleistung von PCs und Workstations ermöglicht es, die spektralen Varianzen direkt durch Fouriertransformation einer Zeitreihe zu schätzen. Dabei ist ein geeignetes Fenster auszuwählen. Die Nichtbeachtung dieser Tatsache führt zwangsläufig zum Rechteckfenster und zum Verschmieren der Information über den ganzen Spektralbereich. Von den hier untersuchten Fenstern hat das Hanning-Fenster die besten Eigenschaften mit dem stärksten Abfall im benachbarten Frequenzbereich.

Analog der Fensterung von Auto- und Kreuzkorrelationsfunktion im Frequenzraum wird das Hanning-Fenster für die Fouriertransformation einer Meßreihe im Frequenzraum angegeben.

Summary

Due to the permanent grow in PC and workstation power it is possible to estimate the spectral variances directly by the Fourier transform of a time series. Thereby one has to choose an appropriate data window. Ignoring this leads to an unsuited square window function where the spectral power in one frequency bin contains leakage from frequency components that are bins away. From the window functions we tested here, the Hanning window has best properties with minimum leakage. Analogous to windowing of the auto- and crosscorrelation function in the frequency domain we give also the Hanning window for the Fourier transform of time series in the frequency domain.

1. Einleitung

Die spektrale Varianzanalyse ist ein fundamentales Werkzeug bei der Datenanalyse in den geophysikalischen Disziplinen, insbesondere in der Meteorologie. Die Bedeutung der Fensterfunktion wird dabei aber meist vernachlässigt. Bei der allgemeinen Verfügbarkeit leistungsstarker Rechner, ergibt sich die Möglichkeit, die Leistungsspektren direkt aus den Daten durch Anwendung von Programmbibliotheken zur Fouriertransformation und nicht über die Auto- und Kreuzkorrelationsfunktion zu gewinnen. Hierbei sind grundlegende Überlegungen in den Hintergrund getreten, wobei nach dem Wiener-Chintschin-Theorem (Taubenheim, 1969, S. 224) für einen Gaußschen Zufallsprozeß beide Betrachtungsweisen äquivalent sind. Bei Press et al. (1992, S. 545 ff.) werden einige Fenster vorgestellt. Bei der Anwendung dieser Fenster ist ihre Wirkung auf reine Sinusschwingungen mit zufälliger Frequenz und Phase entscheidend, denn jede Realisierung einer Zufallsgröße läßt sich aus solchen Schwingungen darstellen.

In der spektralen Varianzanalyse berechnet man den Erwartungswert einer statistischen Größe aus unendlich vielen Realisierungen oder unendlich langen Reihen. Geophysikalische Meßreihen haben prinzipiell eine endliche Länge. Diese endliche Länge der Meßreihe führt bei der spektralen Varianzanalyse dazu, daß die Meßreihe durch ein Fenster betrachtet wird. Außerhalb des Beobachtungszeitraums sind die Werte der untersuchten Größe unbekannt und

werden deshalb zu Null gesetzt. Das Fenster $D(\xi)$ hat somit die Form: $D(\xi)=1$ im Meßintervall und sonst $D(\xi)=0$, wobei ξ eine Ortskoordinate oder die Zeit ist (s.u.).

Die so gewonnene Fouriertransformierte der Meßreihe impliziert eine Periodizität des Prozesses: $M(\xi)=M(\xi+L)$, wenn wir mit M den Meßwert bezeichnen und L die Länge des Datensatzes ist. Diese Periodizität ist in den Daten im allgemeinen nicht vorhanden, Anfangswert und Endwert der Registrierung gehen nicht stetig ineinander über. Dieser „Sprung“ im Datensatz führt zu Spektralwerten auf allen Frequenzen (Wellenlängen), die nur diesen „Sprung“ abbilden, aber nicht wirklich im Prozeß vorhanden sind (‘leakage-Effekt‘). Dies hat dazu geführt, daß Fensterfunktionen benutzt wurden, die die Daten an den Enden (jeweils 10% der Meßreihe) mit einer stetig von eins gegen Null gehenden Funktion multipliziert (vgl. Hamming, 1977, S. 226; Stull, 1988, S. 310). Hiermit erreicht man, daß Stetigkeit bei jeder dieser Meßreihe erreicht wird und der Datenverlust klein gehalten wird.

Im folgenden vergleichen wir die Eigenschaften dieses geglätteten Rechteckfensters mit denen des Hanning-Fensters, des Sinusfensters und des Rechteckfensters (das jeder Anwender benutzt, der keine Fensterfunktion verwendet). Analog der Fensterung von Auto- und Kreuzkorrelationsfunktion im Frequenzraum (Taubenheim, 1969, S. 291, Schönwiese, 1992, S. 180) wird das Hanning-Fenster für die Fouriertransformation einer Meßreihe im Frequenzraum angegeben.

2. Die empirische Bestimmung von Energiespektren

Wir wollen der Einfachheit halber von einer Zeitfunktion und deren Transformation in den Frequenzraum sprechen, für eine ortsabhängige Variable lassen sich diese Überlegungen analog durchführen, wobei z.B. Meßwerte, die sich über die ganze Erde erstrecken, eine Ausnahme bilden, da hier die Periodizität gegeben ist. Das Energiespektrum einer geophysikalischen Größe $M(t)$ ist

$$P'(v) = Am'(v)Am'(v)^* \text{ mit } Am'(v) = \int_{-\infty}^{+\infty} M(t)e^{i2\pi vt} dt, \quad (1)$$

wobei t die Zeit und die v die Frequenz ist. Der Stern steht für den konjugiert komplexen Wert. Die gemessenen Daten sind nur im Meßintervall $\{0,T\}$ vorhanden, somit ist der Erwartungswert des Energiespektrums

$$E\{P'(v)\} = P(v) = \int_{-\infty}^{+\infty} D(t)M(t)e^{i2\pi vt} dt \left(\int_{-\infty}^{+\infty} D(t)M(t)e^{i2\pi vt} dt \right)^*, \quad (2)$$

mit

$$\begin{aligned} D(t) &= 1, \quad \text{für } 0 \leq t \leq T, \\ D(t) &= 0, \quad \text{sonst.} \end{aligned} \quad (3)$$

Dieses Fenster wird als Rechteckfenster bezeichnet.

Gleichung (2) beinhaltet ein Faltungsintegral, daß sich durch die Fouriertransformierte von $D(t)$ und von $M(t)$ ausdrücken läßt. Die Fouriertransformierte von $D(t)$ bezeichnen wir mit $Q(v)$, und die von $M(t)$ ist das Amplitudenspektrum $Am(v)$.

$$E\{P'(v)\} = P(v) = \int_{-\infty}^{+\infty} Q(v-v')Am(v')dv' \left(\int_{-\infty}^{+\infty} Q(v-v')Am(v')dv' \right)^*. \quad (4)$$

Gleichung (4) läßt sich durch Zusammenfassen der Integrale weiter vereinfachen. Der in der Diskussion vorgeschlagene Algorithmus für eine vollständige Analyse mehrerer Zeitreihen, ist in Form von Gleichung (4) am einfachsten zu programmieren. Wir werden hier die Eigenschaften des Rechteckfensters an konkreten Eingangsdaten diskutieren und mit anderen Fenstern vergleichen. Die Fouriertransformierte von $D(t)$, $Q(v)$, ist bei Taubenheim (1969, S. 262) ausführlich diskutiert. Idealerweise wäre $Q(0)=1$ und sonst Null. Dies ergibt sich nur im Fall unendlich langer Meßreihen, die aber real nicht existieren. Deswegen erscheinen die Energiedichten aus den Nachbarfrequenzbereichen im untersuchten Frequenzbereich.

$Q(v)$ hat die Form einer Spaltfunktion $\frac{\sin x}{x}$ und hat neben der endlichen Durchlaßbreite noch

Nebenzipfel. Um diese Nebenzipfel zu dämpfen und die Durchlaßbreite zu optimieren, wendet man andere Filter als das in Gleichung (3) an. Es sei noch mal betont, daß bei der spektralen Varianzanalyse ohne Filterung, automatisch das Fenster nach Gleichung (3) angewendet wird und es somit, je nach Periodizität der Meßdaten, zu einem „Verschmieren“ der Spektralwerte über die angrenzenden oder im schlechtesten Fall über alle Frequenzen kommt, so daß das eigentliche Energiespektrum der Daten nicht mehr erkannt werden kann.

3. Betrachtete Fensterfunktionen

Die im folgenden vorgestellten Fensterfunktionen sind in Abb.1 dargestellt.

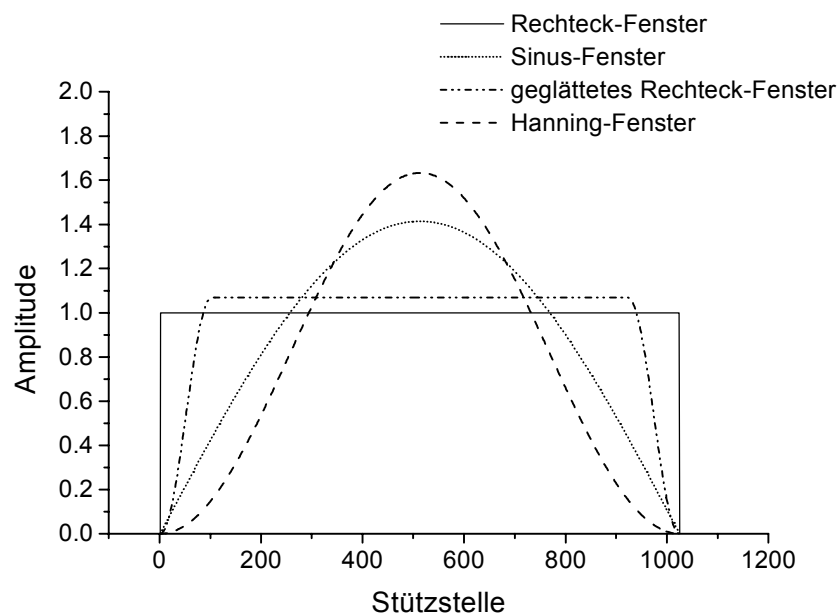


Abb.1 In der spektralen Varianzanalyse untersuchte, normierte (quadratische gemittelte) Fensterfunktionen

Eine oben bereits erwähnte mögliche Fensterform ist die, bei der die Periodizität der Daten durch Multiplikation der Datenreihe an den Enden (jeweils 10% der Meßreihe) mit einer stetig von Null gegen Eins (am Anfang der Datenreihe) bzw. von Eins gegen Null (am Ende der Datenreihe) gehenden Funktion erzwungen wird. Ein solches Fenster kann folgendermaßen gebildet werden (Hamming, 1977, S. 226):

$$\begin{aligned}
D(t) &= \frac{1}{2} \left(1 - \cos\left(\frac{10\pi t}{T}\right)\right) && \text{für } 0 \leq t \leq T/10, \\
D(t) &= 1, && \text{für } T/10 < t < 9T/10, \\
D(t) &= \frac{1}{2} \left(1 + \cos\left(\frac{10\pi(T-t)}{T}\right)\right) && \text{für } T - T/10 \leq t \leq T, \\
D(t) &= 0, && \text{sonst.}
\end{aligned} \tag{5}$$

Wegen der Form nennen wir dieses Fenster geglättetes Rechteckfenster. Ein ähnliches Fenster, aber mit einer anderen Funktion an den Datenenden, findet man bei Stull (1988, S. 310).

Außerdem haben wir ein Sinusfenster

$$\begin{aligned}
D(t) &= \sin\left(\frac{\pi t}{T}\right), && \text{für } 0 \leq t \leq T, \\
D(t) &= 0, && \text{sonst,}
\end{aligned} \tag{6}$$

und ein Hanning-Fenster untersucht

$$\begin{aligned}
D(t) &= \frac{1}{2} \left(1 - \cos\left(\frac{2\pi t}{T}\right)\right), && \text{für } 0 \leq t \leq T, \\
D(t) &= 0, && \text{sonst.}
\end{aligned} \tag{7}$$

Daneben gibt es noch eine Vielzahl weiterer Fenster (vgl. Press et al., 1992, S. 546).

4. Untersuchte synthetische Reihen

Eine stationäre geophysikalische Meßreihe ist in der Regel durch einen stochastischen Gaußschen Zufallsprozeß beschreibbar (s.a. Taubenheim, 1969, S. 200 ff.). Es sind demnach unendlich viele Frequenzen in diesem Prozeß enthalten. Die reale Zeitreihe besitzt neben den kontinuierlichen Spektralbereichen auch Energielücken (z.B. Windspektrum) oder bevorzugte Frequenzen (Tagesgang, Jahresgang) oder auch erhöhte Energiedichten mit vorläufig unbekanntem Frequenzen und Phasenlagen. Darum betrachten wir nicht die Fenstereigenschaften bei diskreten Frequenzen (das sind Vielfache der Nyquist-Frequenz $f_N=1/2\Delta t$, mit Δt als dem Digitalisierungsschritt der Zeitreihe), sondern bei zufälligen Frequenzen und zufälligen Phasenlagen.

Wir stellen uns jetzt die Frage: Wie wird eine Sinusschwingung mit zufälliger Phasenlage im Meßintervall $\{0, T\}$ und beliebiger Frequenz im Spektrum abgebildet? Wie muß das Fenster beschaffen sein, damit die Energie im gesamten Spektrum außerhalb der beiden benachbarten Frequenzen, die die Schwingung repräsentieren, minimal wird?

Die Testfunktion hat folgende Form

$$M(t) = \sin\left(2\pi\left(\frac{i + R_1}{T}t + R_2\right)\right) \quad \text{mit } 0 \leq R_{1,2} \leq 1, \text{ (Zufallszahlen)}, \tag{8}$$

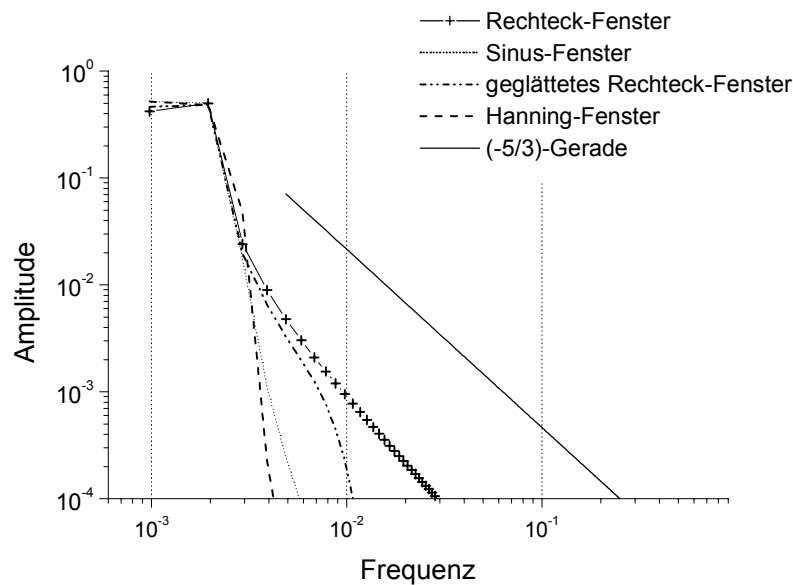
wobei $i < T/2$ und ganzzahlig ist.

Bei jeder Phasenlage liefert ein Fenster einen anderen Filterwert. Deshalb haben wir 100 zufällige Sinus-Reihen mit jeweils gleichem i , aber unterschiedlichem R_1 und R_2 , erzeugt und fouriertransformiert. Das bedeutet, daß die Testfunktion eine reine Sinusschwingung, mit zufälliger Phasenlage und einer zufälligen Frequenz, die auf oder zwischen den beiden Frequenzbins i und $i+1$ liegt, ist. Die Fouriertansformierten wurden nach Gleichung (2) berechnet, aufsummiert und anschließend gemittelt. Auf den Frequenzbins i und $i+1$ wird an-

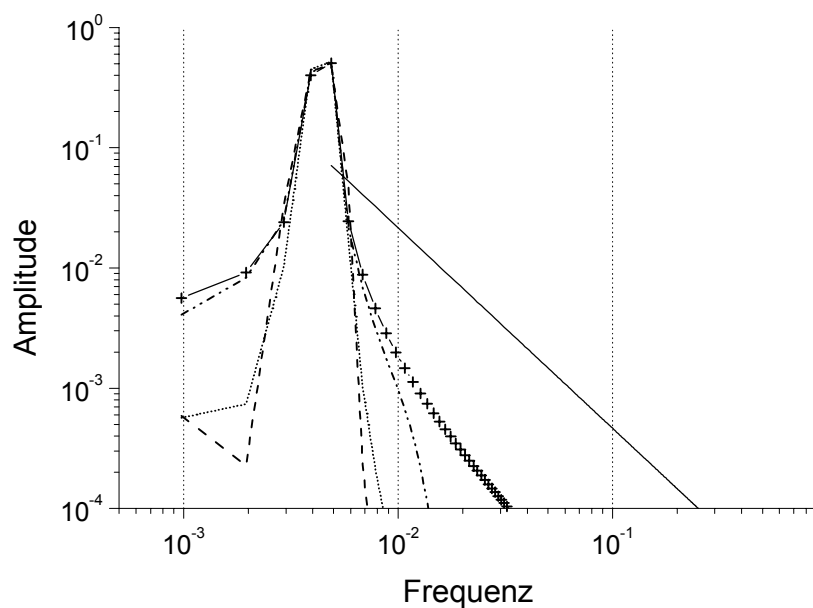
nähernd die gleiche Energie (Leistung) erwartet. Der Abfall der Leistung nach beiden Seiten ist somit ein Maß für die „Güte“ des Fensters.

5. Ergebnisse

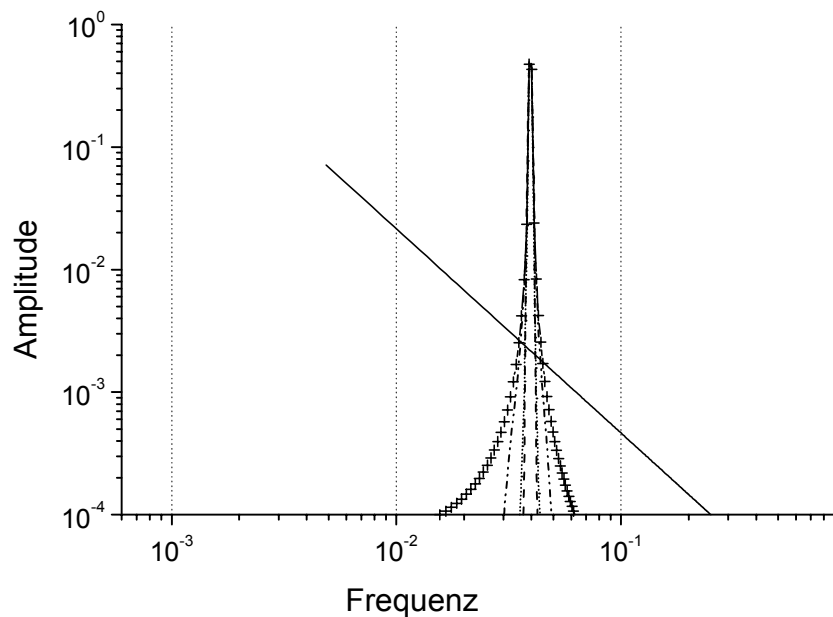
Wir haben eine Fouriertransformation (FFT Routine) benutzt, bei der die Reihenlänge eine Zweierpotenz sein muß, $L=2^N$, mit N einer ganzen Zahl. Bei unseren Rechnungen haben wir $L=1024$ also $N=10$ gewählt. Die hier getroffenen Aussagen beschränken sich nicht auf die FFT, sondern sind allgemeingültig. In den Abbildungen 2a-2c sind die Ergebnisse für $i=1, 4, 40$ dargestellt.



(a) Signal zwischen dem 1. und 2. Frequenzbin



(b) Signal zwischen dem 4. und 5. Frequenzbin



(c) Signal zwischen dem 40 und 41. Frequenzbin von insgesamt 512 Frequenzbins

Abb.2 Energiespektren mit Fensterfunktionen nach den Gleichungen (3) und (5)-(7). Ein Signal, dessen Frequenz zufällig zwischen zwei Frequenzbins liegt, wird auf benachbarten Bins abgebildet. Das Ziel der Fensterung ist es, diese Signale bei großem Offset, wo das Rechteck-Fenster starke Nebenzipfel hat, zu reduzieren. Zum Vergleich ist eine Gerade mit einem Anstieg von $-5/3$ dargestellt.

Wie erwartet ist der Fenstereinfluß bei den niedrigen Frequenzen am größten. Es zeigt sich, daß für diesen Test das Hanning-Fenster eindeutig die besten Eigenschaften besitzt. Das Rechteckfenster bildet niederfrequente Sinusschwingungen auf alle Frequenzen mit einer relativ hohen Energiedichte ab, deren Abfall zu den hohen Frequenzen nur wenig stärker ist als der eingezeichnete $(-5/3)$ -Abfall.

Die Energiedichte geophysikalischer Prozesse ist häufig bei niedrigen Frequenzen am größten und nimmt nach höheren Frequenzen bei einer doppeltlogarithmischen Darstellung häufig nach einem $-5/3$ Gesetz gleichmäßig ab (rotes Rauschen, Turbulenz; Stull, 1988, S. 390 f.). Wir haben in den Abbildungen 2a-2c diese $-5/3$ Linie mit eingezeichnet. Durch den Verzicht auf eine Fensterfunktion sind diese Spektren stark verfälscht und zeigen einen einheitlichen Abfall im hochfrequenten Bereich. Aber gerade dieser Abfall der Energie in diesem Bereich ist häufig Gegenstand der Untersuchung. Beim Sinus-Fenster und auch beim geglätteten Rechteckfenster ist dieser Effekt nicht so stark, wobei das geglättete Rechteckfenster in der Nähe der zu filternden Schwingung kaum besser ist als das Rechteckfenster. Bei größeren Abständen von der zu filternden Frequenz sind die „durchsickernden“ Amplituden sowieso recht klein, so daß das geglättete Rechteckfenster einen größeren Rechenaufwand bereitet, aber wenig Erfolg zeigt.

Betrachten wir den Fall $i=4$, d.h. die Testsinusreihen haben Frequenzen zwischen dem 4. und 5. Frequenzbin (siehe Abb. 2b). Das Hanning-Fenster hat auf dem 3. und 6. Frequenzbin noch die höchsten Energien mit $1/10$ des Maximalwertes, aber beim 2. und 7. Frequenzbin ist die Energie auf drei Größenordnungen abgesunken. Das Rechteckfenster und des geglättete Rechteckfenster liegen hier noch dicht beieinander und haben gerade 1,5 Größenordnungen

erreicht.

Man kann sich die Wirkung der Fenster am einfachsten veranschaulichen, wenn man in Gleichung (8) $i=1$, $R_1=0.5$, $R_2=0$ einsetzt und die so entstehende Reihe periodisch fortführt. Ähnlich wie bei einer Sägezahnfunktion wird die entstandene Unstetigkeit durch hochfrequente Anteile im Spektrum repräsentiert. Die Multiplikation mit den entsprechenden Filtern schwächt diese Unstetigkeit ab. Der Leser wird sofort feststellen, daß im Fall $R_1=0$ aber R_2 beliebig, alle Filter das Ergebnis verschlechtern. Diese Fälle sind in den Abbildungen mit enthalten, sie sind genauso wahrscheinlich wie jede andere Phasenlage auch.

6. Diskussion

Das Hanning-Fenster hat zwei Vorteile: Es unterdrückt die nicht vorhandenen Frequenzen am stärksten, im Vergleich mit den anderen, hier betrachteten Fensterfunktionen und es ist einfach in der Anwendung (s.u.).

Die Multiplikation der Zeitreihe mit der Fensterfunktion erscheint etwas umständlich, darum wollen wir auf die zweite Möglichkeit der Anwendung hinweisen.

Gleichung (4) liefert ein Faltungsintegral im Frequenzraum, das in seiner praktischen Ausführung sehr einfach wird. Die Fouriertransformierte des Hanning-Fensters $Q(v_i)$ hat in der diskreten Form die Werte: $Q(v_{-1})=-0.25$, $Q(v_0)=0.5$, $Q(v_1)=-0.25$ und $Q(v_i)=0$ für alle anderen Frequenzen. Damit ergibt sich die diskrete Rechenvorschrift für das Hanning-Fenster im Frequenzraum (analog dem Faltungsintegral in Gleichung (4) in Summenform) wegen $Am(v_0)=0$ zu

$$Am(v_i) = \frac{1}{4}(-Am'(v_{i-1}) + 2Am'(v_i) - Am'(v_{i+1})),$$

$$Am(v_1) = \frac{1}{4}(2Am'(v_1) - Am'(v_2)), \quad (9)$$

$$Am(v_{\max}) = \frac{1}{4}(-Am'(v_{\max-1}) + 2Am'(v_{\max})).$$

Vergleicht man diese Formel mit der von Taubenheim (1969, S. 291) und Schönwiese (1992, S. 180) angegebenen, so stellt man fest, daß sie sich um ein Vorzeichen unterscheiden. Wir müssen uns dabei vor Augen halten, daß wir über die Zeitreihe ein Hanning-Fenster legen und nicht über die Autokorrelationsfunktion. Das Hanning-Fenster in Gleichung (7) und bei Taubenheim (1969, S. 291) und Schönwiese (1992, S. 179) unterscheiden sich ebenfalls um ein Vorzeichen. Gleichung (9) läßt sich sehr einfach in ein Tabellenkalkulationsprogramm einschließen, wenn das Programm eine Fouriertransformationsroutine anbietet.

Wie bei Schönwiese (1992, S. 183) diskutiert, können bei der Berechnung des Leistungsspektrums aus der Autokorrelationsfunktion numerische Instabilitäten auftreten. Man erkennt sie am Erscheinen negativer Energien. Im Gegensatz dazu bleibt das Spektrum stabil, wenn es direkt aus der Zeitreihe nach Gleichung (4) mit dem diskreten Fenster nach Gleichung (9) berechnet wird. Ursache hierfür ist die zweimalige Anwendung des Fensters, sowohl bei der Fouriertransformierten der Zeitreihe als auch beim konjugiert komplexen, so daß das Ergebnis immer reell und positiv ist. Dies ist auch bei der Normierung der Fenster zu beachten. Die Filter müssen so normiert sein, daß folgende Gleichung erfüllt ist und somit die Varianz nach der Transformation erhalten bleibt:

$$\sigma^2 = \frac{1}{L} \sum_{i=1}^L M(t_i)^2 = \frac{1}{2} \sum_{i=-L/2+1}^{L/2} P(v_i). \quad (10)$$

Die Leistung berechnet sich aus den Effektivwerten der Amplituden (siehe auch Taubenheim, 1969, S. 225), so daß die einzelnen Filterkoeffizienten c_i , die sich nach Gleichungen (5)-(7) berechnen, die Bedingung erfüllen müssen:

$$\sum_{i=1}^L c_i^2 = 1. \quad (11)$$

Damit die hier vorgestellten Fenster Gleichung (11) erfüllen, sind sie mit folgenden Faktoren zu multiplizieren:

Rechteckfenster:	1
geglättetes Rechteckfenster:	1,06875
Sinusfenster:	$\sqrt{2}$
Hanning-Fenster:	$2\sqrt{2/3}$

Die unterschiedlichen Ergebnisse bei der Fouriertransformation mit Hanning-Fenster aus der Autokorrelationsfunktion und aus der direkten Transformation der Meßreihe sind kein Bruch des Wiener-Chintschin-Theorem (Taubenheim, 1969, S. 224) für einen Gaußschen Zufallsprozeß, sondern darauf zurückzuführen, daß das Hanning-Fenster unterschiedlich angewandt wird. Beide Betrachtungsweisen sind zwar äquivalent, aber im ersten Fall wirkt das Hanning-Fenster und im zweiten Fall das Quadrat des Hanning Fensters.

7. Programmalgorithmus zur spektralen Varianzanalyse

Hat man zwei Datensätze und berechnet die Fouriertransformierten, kann man mit Gleichung (4) auch das Kreuzspektrum, die Kohärenz und die Phase berechnen (Stull, 1988, S. 331), indem man den konjugiert komplexen Teil durch das konjugiert komplexe Spektrum der zweiten Reihe ersetzt. Nach einmaliger Durchführung dieser Transformation ist die Varianz des Spektrums auf dem Frequenzbin i überall gleich dem Quadrat des Erwartungswertes auf dieser Frequenz, die Standardabweichung gleich 100% (Press et al., 1992, S. 544-545). Die normierte Kohärenz nach Stull (1988, S. 332) ist nicht zufällig, sondern immer und auf allen Frequenzen gleich eins. Die Fakten bleiben bei Nichtanwendung eines Fensters die gleichen, wobei wir nochmals betonen, das Rechteckfenster wird in diesem Falle immer angewandt.

Die Probleme lassen sich relativ einfach lösen. Wir nehmen an, wir besitzen 10240 äquidistant erfaßte Daten. Diese Meßreihe wird in 10 gleiche Teile von 1024 geteilt und für jede Teilreihe die Fouriertransformation mit dem Hanning-Fenster durchgeführt. Es werden 10 Spektren berechnet. Sind mehrere Reihen vorhanden, werden auch die komplexen Kreuzspektren über diese 10 Teilreihen berechnet. Alle Spektren werden anschließend summiert und durch 10 dividiert (Press et al., 1992, S. 544-545). Das Ergebnis entspricht einem Spektrum, das aus einer Auto(Kreuz-)korrelationsfunktion mit einer maximalen Verschiebung von 10 berechnet wurde, also einem Freiheitsgrad von etwa 20 (siehe Taubenheim, 1969, S. 293). Aufbauend auf diesen Freiheitsgrad lassen sich die entsprechenden Signifikanztests durchführen, die Standardabweichung ist nicht mehr 100%, die aus dem gemittelten Kreuzspektrum und den Leistungsspektren berechnete normierte Kohärenz ist dann in der Regel kleiner als eins, zufällige Phasenlagen löschen sich in den Kreuzspektren aus. Auch hier und bei der berechneten Phase sind so entsprechende Signifikanztests möglich.

Den „Datenverlust“ bei der Filterung mit dem Hanning-Fenster kann man extrem verkleinern, indem man die Reihe in 19 Teile zu 1024 Werten aufteilt, die sich jeweils zu 512 Werte überlappen. Das oben gesagte wird auf diese 19 Teilreihen angewandt. Die heutige Rechenkapazität der PCs ist so gewaltig, daß man den Unterschied in der Rechenzeit kaum bemerkt.

Es sei aber noch darauf hingewiesen, daß sich der Freiheitsgrad durch diese Vorgehensweise nicht erhöht, sondern daß der Freiheitsgrad beim vorhergehenden Algorithmus, abhängig vom gewählten Fenster, eher kleiner als 20 ist. Das hier gewählte Zahlenbeispiel läßt sich sinngemäß auf alle in der Praxis vorkommenden Fälle übertragen.

8. Literatur

Hamming, R.W., 1977: Digital Filters. - Englewood Cliffs, N.J. : Prentice-Hall

Taubenheim, J., 1969: Statistische Auswertung geophysikalischer und meteorologischer Daten. – Leipzig : Akad. Verl.Ges.

Press, W. H., S. A. Teukolsky, W. T. Vetterling und B. P. Flannery, 1992: Numerical Recipes in FORTRAN. The Art of Scientific Computing. - Second Edition. – Cambridge : Cambridge University Press

Schönwiese, Ch., 1992: Praktische Statistik. – 2., verb. Aufl. – Berlin [u.a.] : Borntraeger

Stull, R. B., 1988: An Introduction to Boundary Layer Meteorology. - Dordrecht : Kluwer Academic Publishers

9. Weiterführende Literatur

Harris, F.J., 1978: On the use of windows for harmonic analysis with the discret Fourier transform. Proc. IEEE. 66. S. 51-83

Nuttall, A.H., 1981: Some windows with very good sidelobe behaviour. IEEE Trans. Acoust., Speech, Sign. Proc. 29: S. 84-91

Anschriften der Verfasser:

Dr. H.-J. Schönfeldt, S. von Löwis

LIM - Institut für Meteorologie, Universität Leipzig

Stephanstr. 3, 04103 Leipzig

e-mail: schoenfeldt@uni-leipzig.de

A Possible Connection of Midlatitude Mesosphere/Lower Thermosphere Zonal Winds and the Southern Oscillation

Christoph Jacobi and Dierk Kürschner

Summary:

Collm (52°N, 15°E) midlatitude mesosphere/lower thermosphere (MLT) zonal winds are investigated with respect to a possible influence of the Southern Oscillation (SO). It is found that in winter the prevailing wind is positively correlated with the Southern Oscillation index (SOI), which qualitatively fits to the response of the low-latitude stratosphere on the SO. In summer, however, there is a negative correlation between MLT zonal winds and SOI. A possible mechanism, which includes slightly enhanced planetary wave propagation to the MLT in summers of low SOI is discussed.

Zusammenfassung:

Windmessungen aus Collm (52°N, 15°E) wurden hinsichtlich eines möglichen Einflusses der Southern Oscillation (SO) in ihnen untersucht. Es zeigt sich, dass im Winter der mittlere Zonalwind positiv mit dem "Southern Oscillation Index" SOI korreliert ist, was qualitativ mit der Reaktion der Stratosphäre auf ENSO übereinstimmt. Im Sommer findet sich jedoch ein negativer Zusammenhang zwischen hochatmosphärischem Zonalwind und SOI. Ein möglicher Mechanismus, der die Ausbreitung planetarer Wellen beinhaltet, wird diskutiert.

1. Introduction

The Southern Oscillation (SO), often considered in connection with El Niño and thus abbreviated ENSO, is one of the strongest signals of the tropical atmosphere, which influences also extratropical regions. The SO is a seesaw in surface pressure between the tropical Indian Ocean/eastern Pacific and the other parts of the tropical Pacific, which influences tropospheric temperature and other climate parameters on a global scale. A typical period range of the connected oscillation is 3 - 4 years.

The SO also influences the stratosphere, as has been shown by van Loon and Labitzke (1987). During so-called "warm events" of the SO in winter the Aleutian high is intensified and the stratospheric polar vortex is disturbed. The reason for that is that during warm events the warm Pacific sea surface then leads to stronger convection in the tropics, thus stronger radiation from the cloud tops in the lower stratosphere and a colder tropical stratosphere then (e.g. Labitzke, 1999).

As the SO has some influence upon the stratosphere, it can be assumed that it also influences the upper layers of the middle atmosphere, namely the mesosphere/lower thermosphere (MLT) region around 70 - 110 km altitude, which is dynamically connected with the stratosphere e.g. through planetary and gravity wave propagation. MLT wind variations in the period range of 2 - 5 years were reported by Namboothiri *et al.* (1993, 1994) using the Saskatoon, Canada, medium frequency (MF) radar data. Oscillations with periods of 2 - 5 years were also found within time series of the prevailing wind by Portnyagin (1986) and

Portnyagin *et al.* (1993). Also recently, Kane *et al.* (1999), using an extended database of Saskatoon winds, found variations in the 3-4 years band and speculated that these may be due to an influence of ENSO. Gavrilov *et al.* (1999) have found some interannual variability in gravity wave activity which they claimed to partially being due to ENSO. Investigating time series of MU radar mesospheric winds over Shigaraki, Japan (35°N) they found during times of El Niño weaker westerlies or stronger easterlies in winter and summer, respectively.

To investigate, whether the ENSO signal is also visible in the mesosphere/lower thermosphere (MLT) region we investigate wind data from the long-term measurements over Collm, Germany. This is one of the very few MLT wind time series that extend over more than one decade. One has to be aware that results from one single point measurement can possibly not be assigned to the whole MLT and regional differences may be present. So, for instance, Jacobi *et al.* (2000) showed that there are considerable differences in the year-to-year variation of MLT winds over Canada and Europe at the same latitude. Therefore the results from Collm winds alone can only be attributed to the MLT region over Europe.

2. Data

The horizontal wind field of the mesosphere/lower thermosphere region over Collm, Germany, is continually observed by daily D1 radio wind measurements in the LF range, using the ionospherically reflected sky wave of three commercial radio transmitters on 177, 225 and 270 kHz. The measurements are carried out according to the closely-spaced receiver technique. A modified form of the similar-fade method is used (Schminder and Kürschner, 1994). The procedure is based on the estimation of time differences between corresponding fading extrema of the ionospherically reflected LF sky wave for three measuring points forming a right angled triangle over the ground with small sides of 300 m in direction N and E, respectively. The individual pairs of time differences which allow the calculation of individual wind vectors are measured at a temporal resolution of 0.25 s.

The individual wind measurements are combined to half-hourly zonal and meridional mean wind values on each frequency, with a mean value being averaged over 30 - 60 data points per frequency. The 1- σ variation of the half-hourly mean is in the order of 20 ms⁻¹, caused by the real wind variations and the uncertainties and number of the individual wind measurements. Since 1979, in order to improve the statistical reliability of the results, including the results of the individual measurements on each of the three frequencies combined with a weighting function based on an estimate of the "chaotic velocity" (Sprenger and Schminder, 1969), mean values are calculated, referring to a reflection point at 52°N, 15°E. Since during the daytime the absorption of the sky wave is large, the measuring density decreases strongly then and especially in summer at times no half-hourly means can be constructed during the daytime. This results in a total of about 700 half hourly mean values per month in summer, and about 1200 in winter.

The prevailing wind calculations are based on half-hourly horizontal wind measurements at the mean measured reflection height of low-frequency radio waves. Thus, the measurements use total reflection of the LF radio waves, so that at any time wind measurements are possible only at one specific height. However, since this height changes during the day (e.g. Kürschner *et al.*, 1987) and also from day to day it is possible to calculate monthly mean profiles of wind

parameters between about 85 - 105 km (depending on season) using a regression analysis with quadratically height-dependent parameters, which takes into account the semidiurnal tide:

$$v_z = \sum_{k=0}^p \left\{ h^k a_{k,z} + b_{k,z} h^k \sin \omega t + c_{k,z} h^k \cos \omega t \right\}, \quad (1)$$

$$v_m = \sum_{k=0}^p \left\{ h^k a_{k,m} + b_{k,m} h^k \sin \omega t + c_{k,m} h^k \cos \omega t \right\}, \quad (2)$$

with $p = 2$, ω as the angular frequency of the semidiurnal tide, and h being estimates of the real reflection height, which is calculated from the virtual reflection height using realistic electron density profiles (e.g. Jacobi *et al.*, 1997). The virtual reflection heights are estimated using measured travel time differences between the ground wave and the reflected sky wave by means of phase comparisons in the modulation frequency range on 1.8 kHz of the amplitude modulated LF radio wave (Kürschner *et al.*, 1987).

The reflection height measurements have begun in September 1982 and are, as well as the wind measurements, carried out continuously since. Therefore mean profiles of MLT wind parameters can be obtained since then, whereas the measurements before late 1982 can only be interpreted in terms of the wind field of the maximum measurement density height near 95 km. For this purpose Eqs. 1 and 2 are applied without height dependence, i.e. with $p = 0$ to the monthly median half-hourly wind values. In the following we shall shortly present long-term mean zonal prevailing wind profiles, but in order to extend the time series back to 1979 for trend estimations we shall use the monthly mean winds calculated without height dependence of the regression coefficients.

In order to improve the separation and the spectral selectivity of the evaluation of the tidal components, we make use of the circular polarization of the tide (Kürschner, 1991; Jacobi *et al.*, 1999) to calculate the tidal amplitudes and phases, thus

$$b_{k,m} = -c_{k,z}, \quad \text{and} \quad c_{k,m} = b_{k,z}. \quad (3)$$

The polarisation of the tidal components was proved e.g. by Jacobi *et al.* (1999) comparing tidal components measured independently at several northern midlatitude stations. The zonal prevailing wind is calculated after:

$$v_{oz}(h) = \sum_{k=0}^p h^k a_{k,z}. \quad (4)$$

This procedure is described, for example, by Kürschner and Schminder (1986) and Jacobi *et al.* (1997). Using this method, also combined profiles of winds from the Collm LF D1 measurements and the Juliusruh MF radar of the IAP Kühlungsborn are regularly calculated (e.g. Schminder *et al.*, 1997).

For describing the SO using one single time series the Southern Oscillation Index (SOI) is frequently used. There exist several slightly different methods in calculating the SOI. It is taken here as the difference of the normalised Darwin and the normalised Tahiti sea level pressure anomaly (Trenberth, 1984; Ropelewski and Jones, 1987). Note that low SOI values mean strong pressure differences across the Pacific, weaker trade-winds and thus warmer sea-surface temperatures over the eastern Pacific. El Niño events are connected with low SOI.

3. Results

3.1 Zonal prevailing winds over Collm

In Figure 1 winds from the HWM93 empirical wind model (Hedin *et al.*, 1996) are shown for winter (December - February, DJF) and summer (June - August, JJA) for 52°N latitude. In the MLT region the respective Collm zonal winds are added. The Collm data are 1983 - 1998 mean values.

Height measurements at Collm are only available since 1983. To obtain an extended time interval, in the following winds calculated without height dependence are used, which can be attributed to the height of the maximum measuring density near 95 km. In Figure 2 the times series for JJA and DJF means are shown together with the respective SOI values. Comparing the Collm winds with the Shigaraki MU radar data (Gavrilov *et al.*, 1999, their Figure 1) it can be seen that, despite the longitudinal and latitudinal difference, some similarities are found, particularly the 1992 minimum in winter is found in both time series, and in the 1990s the Collm summer westerly winds are increasing, which corresponds to the decreasing of the easterly winds over Shigaraki after 1992.

The correspondence between SOI and MLT zonal winds is visible in the winter time series, e.g., in the corresponding minima in 1983, 1992, and 1998. The summer time series appear to be less strongly correlated, but a weak anticorrelation may be visible in the upper part of Figure 2. In addition, a positive long-term trend is visible in the MLT zonal winds with an increase of the westerly winds of $0.51 \text{ ms}^{-1}\text{yr}^{-1}$ in summer and $0.29 \text{ ms}^{-1}\text{yr}^{-1}$ in winter, which is not found in the SOI time series.

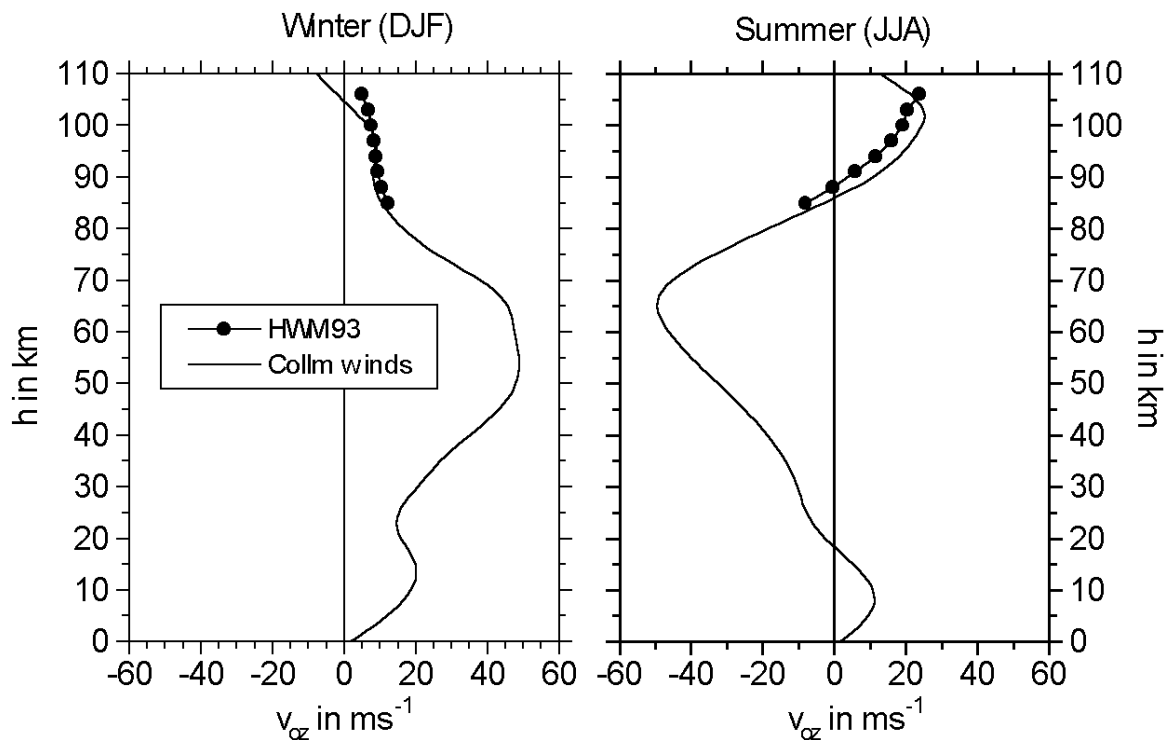


Figure 1: Mean winter (left panel) and summer (right panel) zonal prevailing wind profiles for 52°N from the HWM93 model and from LF D1 measurements over Collm.

3.2 The connection between SO and MLT zonal winds

To show the connection between v_{oz} and SOI, the correlation coefficients between these parameters are shown in Figure 3 for each month of the year. A clear seasonal cycle is visible, although the correlation is not very strong in most months. However, at least during most of the winter the correlation is significant (t-test, see dotted lines in Figure 3).

It has been shown by several authors, and this is also confirmed by the results presented in Figure 2, that the MLT wind parameters exhibit considerable long-term trends and decadal variations, which are possibly owing to long-term atmospheric variability in connection with greenhouse gas emissions and extraterrestrial forcing in the frame of the 11-year solar cycle (e.g. Namboothiri *et al.*, 1993; Bremer *et al.*, 1997; Jacobi *et al.*, 2000). As these signals not necessarily are found in the SO, we apply a high-pass filter by investigating the year-to-year differences of the MLT zonal winds and the SOI. The respective correlation coefficients for these data are also shown as open symbols in Figure 3. It can be seen that in summer the correlation is stronger than if one uses the unfiltered time series. In addition, we also correlated the SOI values of the respective months with the v_{oz} values one month later (crossed symbols). This appears to be reasonable if we may assume that in summer planetary wave (PW) propagation during different phases of the SO may play a role in the possible troposphere-mesosphere-lower thermosphere coupling process, and the waves may need some time to propagate to the lower thermosphere. The correlation is much stronger now for June and July. In winter there is no strong difference when using either filtered or unfiltered time series, because the winter v_{oz} long-term trend is weaker than the summer one so that the filter will only weakly influence the overall variability.

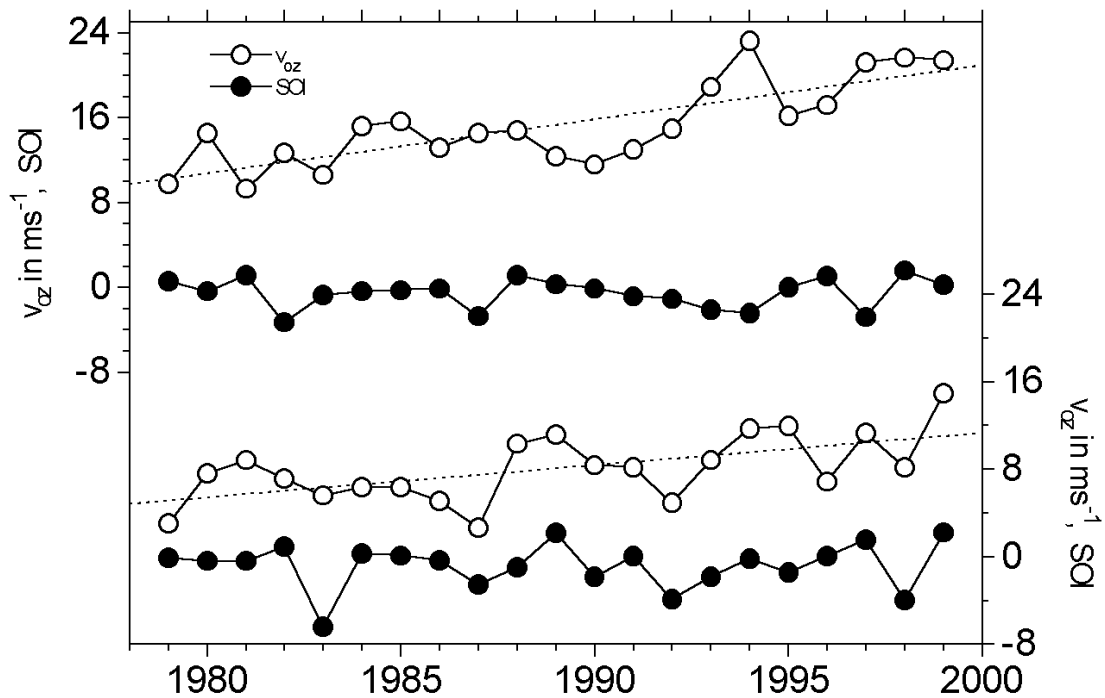


Figure 2: Time series of summer (JJA, upper part of figure) and winter (DJF, lower part of figure) zonal prevailing winds v_{oz} near 95 km and SOI.

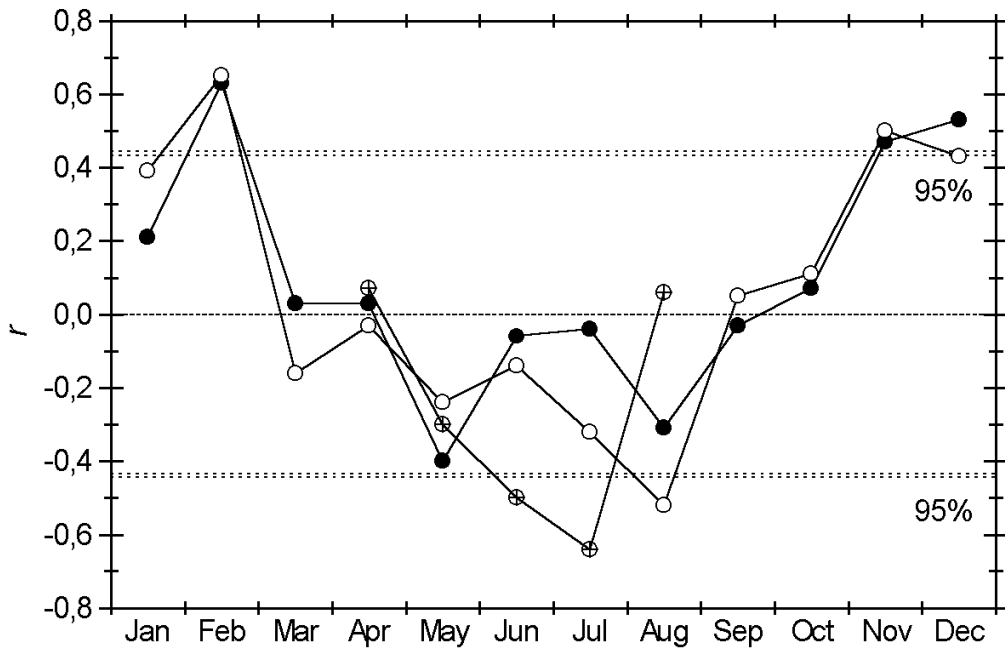


Figure 3: Solid symbols: correlation coefficients r between monthly mean Collm zonal prevailing winds v_{oz} and the SOI for each month of the year. Open symbols: correlation coefficients between the year-to-year differences of SOI and v_{oz} in the respective months. Crossed symbols: correlation coefficients between the year-to-year differences of SOI in the respective months and v_{oz} one month later. Period used is 1979 - 1999. The 95% significance level is indicated by the dotted lines (t-test).

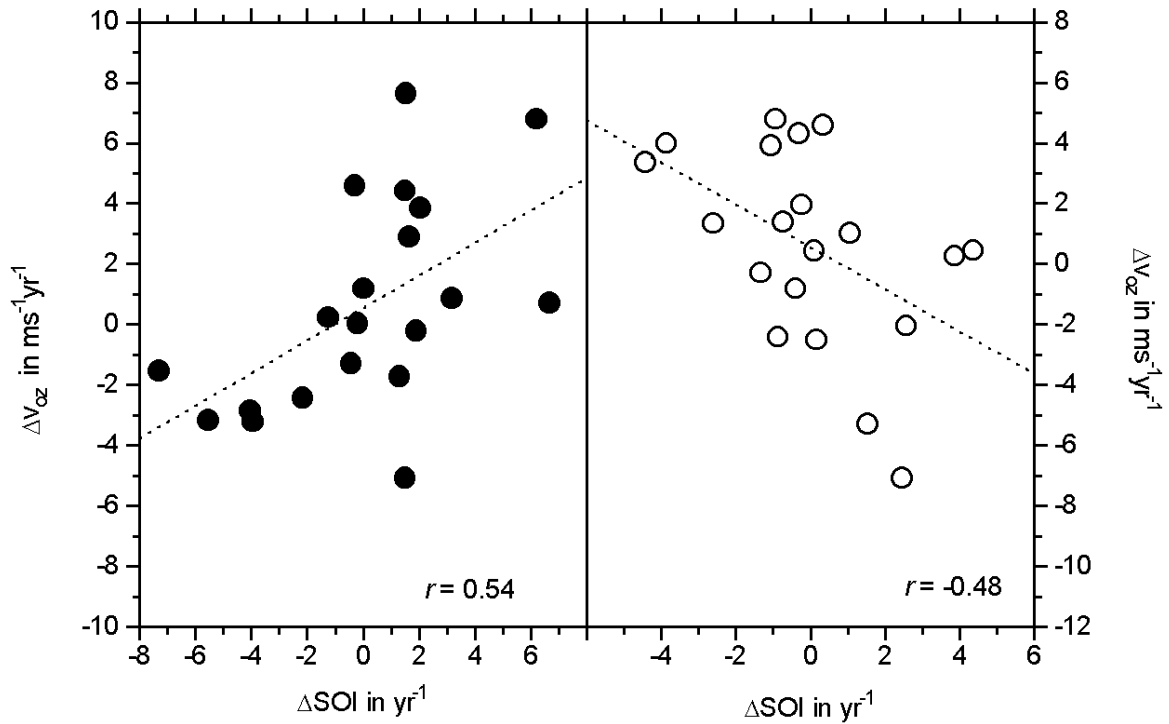


Figure 4: Year-to-year difference of zonal prevailing wind in winter (DJF, left panel) and summer (JJA, right panel) vs. year-to-year difference of SOI.

In Figure 4 these year-to-year differences of the Collm MLT v_{oz} (Δv_{oz}) are presented vs. the respective SOI differences (ΔSOI) for winter (DJF) and summer (JJA). The correlation is positive in winter, which can be expected if the reaction of the stratosphere on warm events (corresponding to small SOI) is taken into account: the stratospheric vortex is weaker due to an intensified Aleutian high (van Loon and Labitzke, 1987), which leads to weaker westerly winds in the mesosphere and also lower thermosphere. Therefore a positive correlation of SOI and v_{oz} is expected. In summer, however, the correlation is negative, although it is not too strong ($r = -0.48$). However, if we correlate May - July mean values of SOI with June - August means of v_{oz} to take into account a possible delay due to the travel time of wave propagation, the correlation strengthens ($r = -0.53$, not shown in the figure).

3.3 Planetary wave propagation and SO in summer

The possible negative correspondence of summer SOI and MLT zonal winds seen in the right panel of Figure 4 cannot be explained by a simple decrease of the pressure level height gradient between high and equatorial latitudes during cold events (high SOI values). This would lead to a weaker strato-mesospheric easterly jet and in the MLT, i.e. above the wind reversal (Figure 1) to a stronger westerly wind, so that a positive connection would be the result. In addition, we compared the zonal mean summer 30 hPa temperatures at 15°N near the equator with the SOI using year-to-year differences to avoid the influence of the negative stratospheric temperature trend (e.g. Keckhut *et al.*, 1999). We found a weak negative correlation ($r = -0.34, -0.56, -0.38$ for June, July and August, respectively) which would mean that during cold events (high SOI) the low-latitude zonal mean stratospheric temperature is lower and thus the latitudinal pressure gradient should be even stronger, if there is any connection.

We assume, however, that the negative connection seen in Figure 3 could be owing to varying PW propagation during the different phases of the SO. If we hypothesise that during cold events either westward travelling PW are more effectively forced or the propagation of part of the westward travelling waves through the stratosphere is facilitated, in the mesosphere and lower thermosphere this would lead to a westward directed acceleration of the mean flow through wave-mean flow interaction. This would mean that in the MLT region the easterlies (below 90 km altitude, see Figure 1) are stronger or the westerlies above 90 km altitude are weaker, as is the case over Collm (Figure 4). In principle, PW with comparatively short periods as the Rossby quasi 5-day wave, which have higher phase velocity, are more easily able to tunnel through the strato-/mesospheric easterlies in summer. However, a considerable part of the summer PW with short periods are due to atmospheric instability of the mesosphere, as is the case with the quasi 2-day wave (see Plumb, 1983). Meyer and Forbes (1997) proposed that a 6.5-day westward propagating wave is owing to atmospheric instability, too. Therefore investigating short-period (less than 10 days) PW might be misleading, and we consider the quasi 16-day Rossby PW here.

A measure for the activity of the quasi-16-day wave is the standard deviation σ_{12-25} of the *daily* mean winds at Collm. Daily zonal and meridional mean wind values are calculated similar to the monthly means, but using half-hourly winds of only one day. The time series being filtered in the period range 12 - 25 days using a Lanczos filter (see, e.g., Thomson and Chow, 1980) with 100 weights. From the filtered time series daily values of the standard deviations $\sigma_{z,12-25}$ and $\sigma_{m,12-25}$ of the zonal and meridional prevailing wind are calculated using a 32-day time interval each, and the respective values are attributed to the centre of the interval.

From these data 3-monthly means can be calculated. The procedure is described in Jacobi *et al.* (1998). The total standard deviation is calculated after:

$$\sigma_{12-25} = \sqrt{\sigma_{z,12-25}^2 + \sigma_{m,12-25}^2} , \quad (5)$$

JJA time series of σ_{12-25} and SOI are shown in Figure 5. No clear connection can be seen, which is probably due to two reasons.

Firstly, during the westerly phase of the equatorial quasi-biennial oscillation (QBO) westward travelling PW may propagate through the equatorial waveguide to the summer lower thermosphere (Dickinson, 1968). Jacobi *et al.* (1998) have found the QBO signal in the PW activity of the summer MLT over Collm. This wave activity may be stronger than the relatively weak signal due to the SO and may cover the correspondence. Therefore, to clearly show the possible process, only QBO east years should be selected. These are marked by the solid symbols in the SOI time series in Figure 5.

Secondly, volcanic eruptions like the ones of El Chichon in 1982 and the Pinatubo in 1991 lead to a strong warming on the lower stratosphere due to increased absorption then (Angell, 1993; Labitzke *et al.*, 1997). Therefore the SO signal in the stratosphere is not found in the year after the eruptions, as was reported by Labitzke and van Loon (1989). The years of the eruptions are marked by big solid symbols in the PW activity time series in Figure 5.

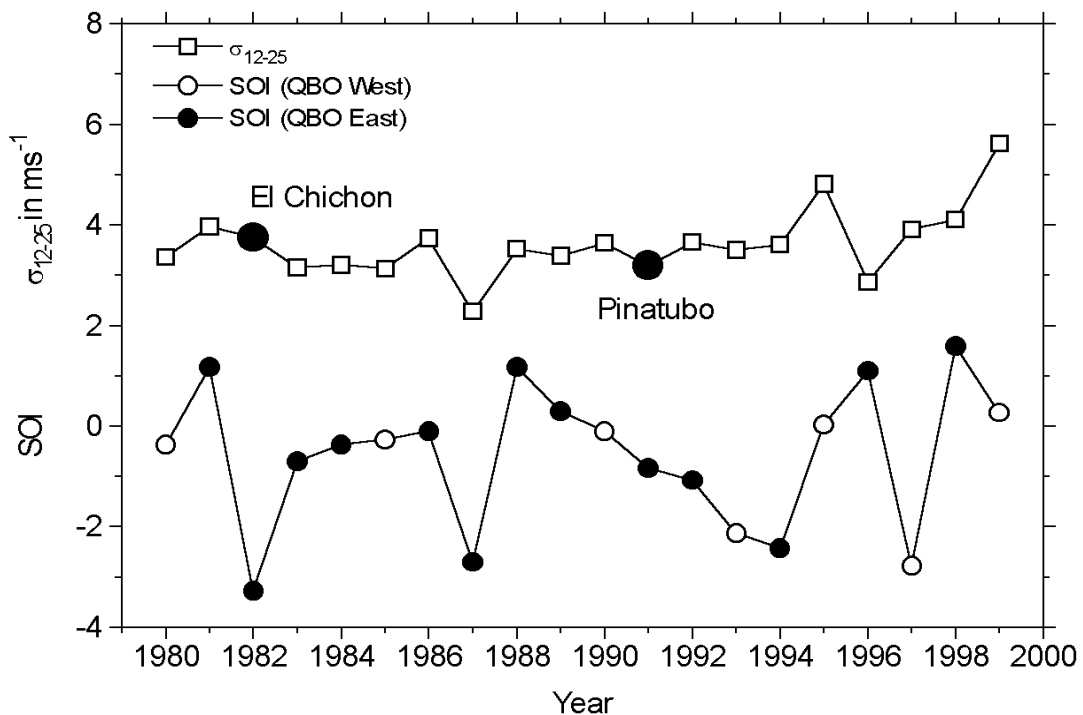


Figure 5: Time series of summer (JJA) prevailing wind standard deviation σ_{12-25} in the 12-25 days range and of the SOI. QBO east years are indicated by a solid symbol of SOI. Thick solid symbols for σ_{12-25} show years that are probably influenced by volcanic eruptions.

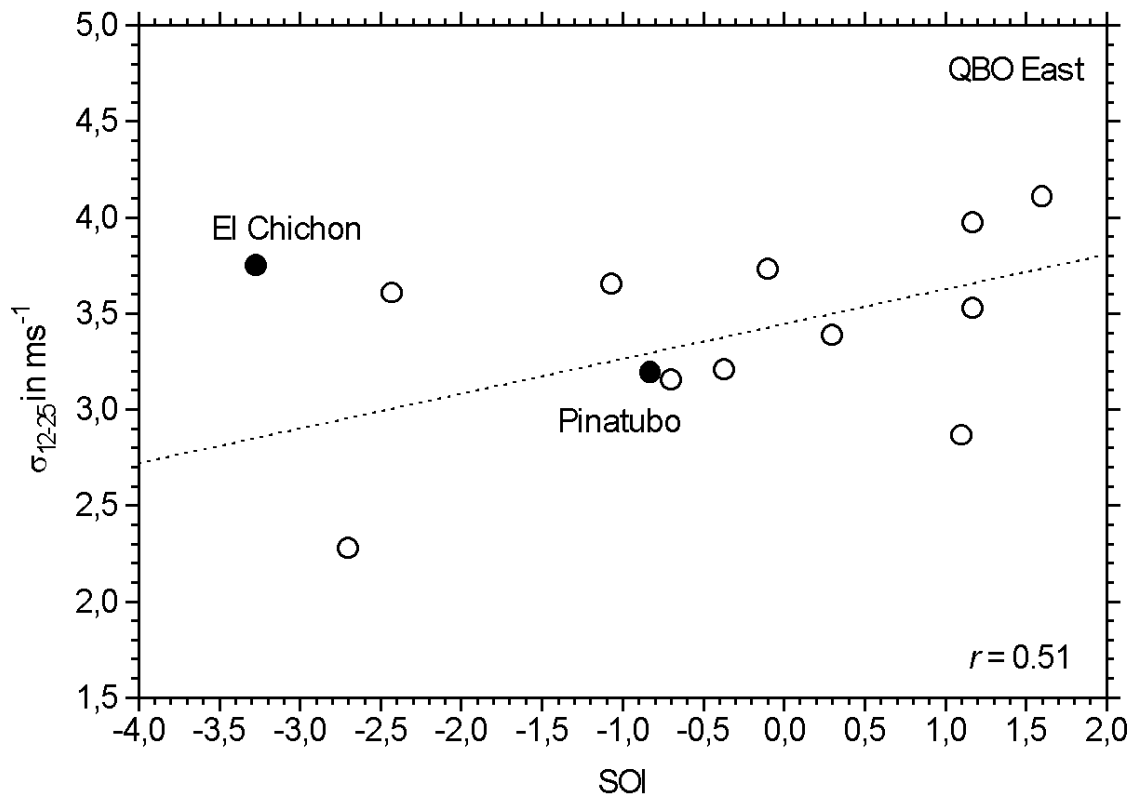


Figure 6: 12-25 day standard deviation of Collm JJA prevailing wind vs. SOI for QBO east years. Solid symbols are years probably influenced by volcanic eruptions.

If we present the σ_{12-25} and SOI date for QBO east years in a scatter plot, as it is done in Figure 6, we find an indication for a correlation between these parameters so that, as expected, the PW activity is slightly enhanced during high SOI years. This is also the case with the stratospheric planetary waves. In the upper panel of Figure 7 the correlation coefficients r between the monthly mean amplitude of the stratospheric height wave of zonal wavenumber 1 and the SOI is presented. The 3-monthly mean wave amplitudes were calculated from daily regression analyses of 30 hPa gridded FU Berlin stratospheric pressure level height data at the respective latitudes. In the lower panel of Figure 7 the 1980 - 1999 mean amplitudes are shown. It can be seen that the correlation coefficients are positive at those latitudes where the amplitudes are strongest, which is in accordance with the Collm MLT results in Figure 6.

It should be noted that the stratospheric data are calculated from radiosonde ascents performed once a day, so that the data may contain the information of the diurnal tide in the stratosphere. However, diurnal tidal amplitudes are maximum at lower latitudes (e.g. Forbes, 1982; Khattatov *et al.*, 1997; Geller *et al.*, 1997; Hagan *et al.*, 1999) while the maximum mean amplitudes in Figure 7 are found at high latitudes. Therefore we conclude that at least a considerable part of the height wave 1 is owing to PW activity.

4. Conclusions

Investigation of monthly Collm MLT zonal winds and SOI show that there is a relationship between the SO and the MLT region, which in winter corresponds to the reaction of the stratosphere to the SO (van Loon and Labitzke, 1987), so that stronger westerly winds are

found during years of high SOI. This is also in correspondence with results from the Japanese MU radar that has been presented by Gavrilov *et al.* (1999).

In summer, Collm MLT zonal winds are negatively correlated to the SOI, while the strongest correlation is found in connection with a phase delay of one month between the SOI and the MLT winds. The connection between SOI and summer MLT winds may be explained by an influence of the SO on the PW propagation conditions in the middle atmosphere, which again force the MLT winds through wave-mean flow interaction. However, the Collm results for the summer months are not in correspondence with the Japanese radar results by Gavrilov *et al.* (1999). They have found stronger easterly winds during a period of El Niño. However, if we detrend the time series of Collm summer winds in the upper panel of Figure 2, we find, as is the case with the MU radar winds (Gavrilov *et al.*, their Figure 1) that there is a general tendency towards weaker thermospheric westerlies (which correspond to stronger mesospheric easterlies) in the early 1990s over both stations. This general tendency is filtered out when regarding only year-to-year variations as done in Figure 4. So possibly there are further processes acting on different time scales that influence the MLT region.

We also have to be aware that the SO in the summer stratosphere is only a weak signal compared, e.g., to the QBO that is the strongest circulation pattern in the tropical stratosphere (e.g. Naujokat, 1986). Therefore PW propagation in the equatorial stratosphere is much more strongly influenced by the QBO than by the SO. Even if we select cases with one particular QBO phase as we did in Figure 7, this does not totally exclude the influence of the QBO. Therefore it is clear that any SO influence on the MLT region must be weak.

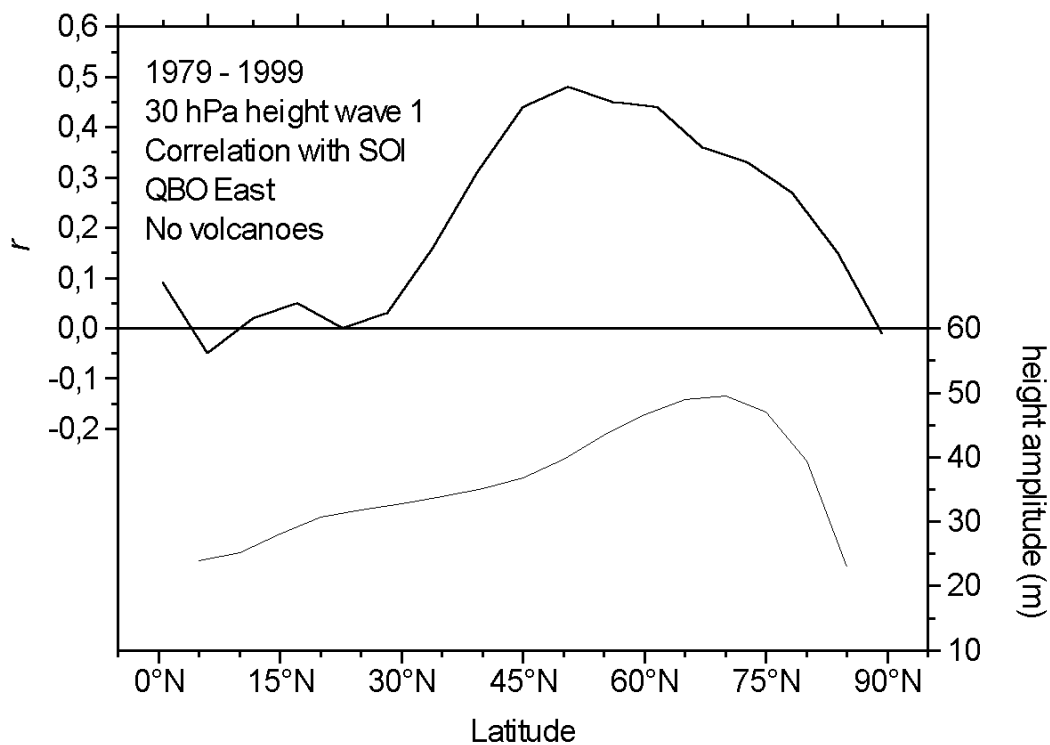


Figure 7: Correlation coefficients r between JJA 3-monthly mean 30 hPa height wave 1 amplitudes and the SOI (upper panel) and monthly mean height wave 1 (lower panel) vs. latitude.

We stress again that results from one single point measurement not necessarily are valid for the entire MLT region and regional differences may be present. Jacobi *et al.* (2000) have presented considerable differences in the year-to-year variation of MLT winds over Canada and Europe. Therefore the results from Collm winds alone can only be attributed to the MLT region over Europe. An extended analysis including a larger number of stations with wind measurements will have to be performed.

Acknowledgements

SOI data were contributed by the NCAR Climate Analysis Section through Internet access on ftp://ftp.cgd.ucar.edu/pub/shear/soi_dir/soi, 9 May 2000. Stratospheric pressure level height data were kindly contributed by the Stratospheric Research Group of the FU Berlin.

References

- Angell, J.K., 1993: Comparison of stratospheric warming following Agung, El Chichon and Pinatubo volcanic eruptions. *Geophys. Res. Lett.* **20**, 715 - 718.
- Forbes, J.M., 1982: Atmospheric tides 1. Model description and results for the solar diurnal component. *J. Geophys. Res.* **87**, 5222 - 5240.
- Gavrilov, N.M., S. Fukao, and T. Nakamura, 1999: Peculiarities of interannual changes in the mean wind and gravity wave characteristics in the mesosphere over Shigaraki, Japan. *Geophys. Res. Lett.* **26**, 2457 - 2460.
- Geller, M.A., V.A. Yudin, B.V. Khattatov, and M.E. Hagan, 1997: Modeling the diurnal tide with dissipation derived from UARS/HRDI measurements. *Ann Geophysicae* **15**, 1198-1204
- Hagan, M. E., M. D. Burrage, J. M. Forbes, J. Hackney, W. J. Randel, and X. Zhang, 1999: GSWM-98: Results for migrating solar tides. *J. Geophys. Res.* **104**, 6813-6828.
- Hedin, A.E., E.L. Fleming, A.H. Manson, F.J. Schmidlin, S.K. Avery, R.R. Clark, S.J. Franke, G.J. Fraser, T. Tsuda, F. Vial, and R.A. Vincent, 1996: Empirical wind model for the upper, middle and lower atmosphere. *J. Atmos. Terr. Phys.* **58**, 1421 - 1447.
- Jacobi, Ch., R. Schminder, and D. Kürschner, 1997: Measurements of mesopause region winds over Central Europe from 1983 through 1995 at Collm, Germany. *Contrib. Atmos. Phys.* **70**, 189 - 200.
- Jacobi, Ch., R. Schminder, und D. Kürschner, 1998: Long-period (12-25 days) oscillations in the summer mesopause region as measured at Collm (52°N, 15°E) and their dependence on the equatorial quasi-biennial oscillation. *Contrib. Atmos. Phys.* **71**, 461 - 464.
- Jacobi, Ch., Yu.I. Portnyagin, T.V. Solovjova, P. Hoffmann, W. Singer, A.N. Fahrutdinova, R.A. Ishmuratov, A.G. Beard, N.J. Mitchell, H.G. Muller, R. Schminder, D. Kürschner, A.H. Manson, und C.E. Meek, 1999: Climatology of the semidiurnal tide at 52°N - 56°N from ground-based radar wind measurements 1985 - 1995. *J. Atmos. Solar-Terr. Phys.* **61**, 975 - 991.
- Jacobi, Ch., M. Lange, D. Kürschner, A.H. Manson, and C.E. Meek, 2000: A long-term comparison of Saskatoon MF radar and Collm LF D1 mesosphere-lower thermosphere wind

- measurements. IAGA Workshop "Lower atmosphere effects on the ionosphere and upper atmosphere"; Prague, 25 - 27.7.2000, submitted to *Phys. Chem. Earth*.
- Kane, R.P., C.E. Meek, and A.H. Manson, 1999: Quasi-biennial and higher-period oscillations in the mean winds in the mesosphere and lower thermosphere over Saskatoon, 52°N, 107°W. *J. Geophys. Res.* **104**, 2645 - 2652.
- Keckhut, P., F.J. Schmidlin, A. Hauchecorne, and M.L. Chanin, 1999: Stratospheric and mesospheric cooling trend estimates from U.S. rocketsondes at low latitude stations (6°S-34°N), taking into account instrumental changes and natural variability. *J. Atmos. Solar-Terr. Phys.* **61**, 447 - 459.
- Khattatov, B.V., M.A. Geller, and V.A. Yubin, 1997: Diurnal migrating tide as seen by the high-resolution Doppler imager / UARS. 2. Monthly mean global zonal and vertical velocities, pressure, temperature, and infrared dissipation. *J. Geophys. Res.* **102**, 4423 - 4435.
- Kürschner, D., and R. Schminder, 1986: High-atmosphere wind profiles for altitudes between 90 and 110 km obtained from D1 FL wind measurements over Central Europe in 1983/1984. *J. Atmos. Terr. Phys.* **48**, 447 - 453.
- Kürschner, D., R. Schminder, W. Singer, and J. Bremer, 1987: Ein neues Verfahren zur Realisierung absoluter Reflexionshöhenmessungen an Raumwellen amplitudenmodulierter Rundfunksender bei Schrägeinfall im Langwellenbereich als Hilfsmittel zur Ableitung von Windprofilen in der oberen Mesopausenregion. *Z. Meteorol.* **37**, 322 - 332.
- Labitzke, K., 1999: Die Stratosphäre. Springer, Berlin, 177p.
- Labitzke, K., and H. van Loon, 1989: The Southern Oscillation. Part IX: the influence of volcanic eruptions on the Southern Oscillation in the stratosphere. *J. Clim.* **2**, 1223 - 1226.
- Labitzke, K., S. Leder, and H. van Loon, 1997: The effect on the lower stratosphere of three volcanic eruptions. NCAR Tech. Note 433+PROC, Boulder, Colorado, 153 - 165.
- Meyer, C.K., and J.M. Forbes, 1997: A 6.5-day westward propagating planetary wave: Origin and characteristics. *J. Geophys. Res.* **102**, 26173 - 26178.
- Namboothiri, S.P., A.H. Manson and C.E. Meek, 1993: Variations of mean winds and tides in the upper middle atmosphere over a solar cycle, Saskatoon, Canada, 52####N, 107####W. *J. Atmos. Terr. Phys.* **55**, 1325 - 1334.
- Namboothiri, S.P., C.E. Meek and A.H. Manson, 1994: Variations of mean winds and solar tides in the mesosphere and lower thermosphere over time scales ranging from 6 months to 11 yr: Saskatoon, 52####N, 107####W. *J. Atmos. Terr. Phys.* **56**, 1313 - 1325.
- Naujokat, B., 1986: An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. *J. Atmos. Sci.* **43**, 1873 - 1877.
- Plumb R.A., 1983: Baroclinic instability of the summer mesosphere: A mechanism for the quasi-two-day wave? *J. Atmos. Sci.* **40**, 262 - 270.
- Portnyagin, Yu.I., 1986: The climatic wind regime in the lower thermosphere from meteor radar observations. *J. Atmos. Terr. Phys.* **48**, 1099 - 1109.

- Portnyagin, Yu.I., J.M. Forbes, G.J. Fraser, R.A. Vincent, S.K. Avery, I.A. Lysenko and N.A. Makarov*, 1993: Dynamics of the Antarctic and Arctic mesosphere and lower thermosphere regions - I. The prevailing wind. *J. Atmos. Terr. Phys.* **55**, 827 - 841.
- Ropelewski, C.F., and P.D. Jones*, 1987: An extension of the Tahiti-Darwin Southern Oscillation Index. *Mon. Wea. Rev.* **115**, 2161 - 2165.
- Schminder, R. and D. Kürschner*, 1994: Permanent monitoring of the upper mesosphere and lower thermosphere wind fields (prevailing and semidiurnal tidal components) obtained from LF D1 measurements in 1991 at the Collm Geophysical Observatory. *J. Atmos. Terr. Phys.* **56**, 1263 - 1269.
- Schminder, R., Ch. Jacobi, D. Kürschner, P. Hoffmann, D. Keuer, and W. Singer*, 1997: The upper mesosphere and lower thermosphere wind field over Central Europe from 1994 through 1996 obtained from a joint analysis of LF windprofiler and MF radar measurements. *Meteorol. Z., N.F.* **6**, 225 - 229.
- Sprenger, K. and R. Schminder*, 1969: On some relationships between the correlation analysis and similar-fade analysis results of drift measurements in the lower ionosphere. *J. Atmos. Terr. Phys.* **31**, 1085 - 1098.
- Thomson, R.E., and K.Y. Chow*, 1980: Butterworth and Lanczos-window cosine filters with application to data processing on the UNIVAC 1106 computer. Institute of Ocean Sciences, Sidney, B.C., Pacific Marine Science Report 80-9.
- Trenberth, K.E.*, 1984: Signal versus noise in the Southern Oscillation. *Mon. Wea. Rev.* **112**, 326 - 332.
- Van Loon, H., and K. Labitzke*, 1987: The Southern Oscillation. Part V: The anomalies in the lower stratosphere of the northern hemisphere in winter and a comparison with the quasi-biennial oscillation. *Mon Wea Rev.* **115**, 357 -369.

Addresses of Authors

Ch. Jacobi, Institut für Meteorologie, Universität Leipzig, Stephanstr. 3, 04103 Leipzig

D. Kürschner, Institut für Geophysik und Geologie, Universität Leipzig, Observatorium Collm, 04779 Wermsdorf

Einfluß des winterlichen stratosphärischen Polarwirbels auf die zonale Symmetrie des Windfeldes in der oberen Mesosphäre und unteren Thermosphäre simuliert mit dem COMMA-Modell

M. Lange, Ch. Jacobi

Zusammenfassung

Langzeitmessungen des Windfeldes in der Mesopausenregion ($\sim 92 \text{ km}$) an verschiedenen Stationen in den mittleren Breiten der Nordhemisphäre zeigen systematische zonale Variationen beim (zeitlich) gemittelten Zonal- und Meridionalwind und bei den Amplituden und Phasen der halbtägigen Gezeiten. Als eines der herausragenden Muster, die zonale Variationen in der unteren mittleren Atmosphäre anregen, wird der Einfluß der Geopotentialstörungen zur zonalen Wellenzahl 1 und 2, die mit dem winterlichen stratosphärischen Polarwirbel verbunden sind, auf das Windfeld in der oberen Mesosphäre / unteren Thermosphäre numerisch mit dem COMMA-Modell der mittleren Atmosphäre untersucht. Die Modellergebnisse zeigen eine gute Übereinstimmung der zonalen Variationen des mittleren Zonalwindes, die im Breitenbereich $52^\circ N$ bis $56^\circ N$ beobachtet werden und in der Größenordnung von $10 - 20 \text{ m/s}$ liegen. Auch die halbtägigen Gezeitenamplituden und -phasen zeigen qualitative und quantitative Übereinstimmungen zwischen Beobachtungen und Modellergebnissen.

Abstract

Long-term time series of wind field observations in the upper mesosphere / lower thermosphere region at different locations in the midlatitude region indicate longitudinal variability in the (time-) mean zonal and meridional wind and in the amplitudes and phases of the semidiurnal tide, too. Being one of the prominent patterns forcing zonal inhomogeneities in the lower middle atmosphere, the influence of the zonal wavenumber 1 and wavenumber 2 disturbances connected with the winter Northern Hemisphere stratospheric polar vortex on the mesosphere- / lower thermosphere wind field is numerically investigated with the COMMA model. The model results show that the zonal variations through the stationary waves coincide with typical observed mean zonal wind differences between different stations along the midlatitude belt between $52^\circ N$ and $56^\circ N$ with values about $10 - 20 \text{ m/s}$. Also, the amplitude and phase variations of the semidiurnal tide show qualitative and quantitative agreements between model results and observations.

1. Einleitung

In numerischen Untersuchungen werden Trends und Wellenausbreitung in der mittleren Atmosphäre häufig durch die zonal gemittelten Felder beschrieben. Langzeitbeobachtungen des Windfeldes in der Mesopausenregion in den mittleren Breiten der Nordhemisphäre

zeigen neben zeitlich und örtlich periodischen Schwankungen durch Gezeiten und planetare Wellen auch systematische längenabhängige Variationen des mittleren Zonalwindes und der Amplitudenverteilung der halbtägigen Gezeiten (Bremer et al., 1997; Jacobi et al., 1999). Als Ursache dafür kommen in erster Linie orografisch bedingte Einflüsse in Frage, die einerseits zu einer ausgeprägten Abweichung des winterlichen Polarwirbels von der Radialsymmetrie um den Pol führen, aber auch zu einer Längenabhängigkeit der Schwerewellenaktivität. Auch Unterschiede in der Land-Meer Verteilung beeinflussen die Schwerewellenaktivität. Weiterhin können direkte Variationen des solaren Antriebs für die halbtägige Gezeitenamplitude durch längenabhängige Ozonvariationen von Bedeutung sein.

Bei dem Vergleich von Langzeitmessungen des halbtägigen solaren Gezeitenwindes in 92 km Höhe an 6 verschiedenen Stationen von Jacobi et. al (1999) (Abb. 1, Tab. 1) wird im Winter eine besonders starke Variation der Amplitude zwischen ca. 8 m/s und 25 m/s und der Phase um ca. 2 Stunden beobachtet. Die Differenz der Amplitude wird mit zunehmendem Abstand zwischen den Stationen größer, was für eine Wechselwirkung der solaren Gezeiten mit einer großskaligen nichtwandernden Wellenstörung spricht. Lediglich die am weitesten entfernte Station (Kazan) weicht davon ab mit der zu Saskatoon vergleichbaren Amplitudenstärke. Auch zeigen Vergleiche zwischen Collm und Saskatoon eine Korrelation der Differenzen der halbtägigen Gezeitenamplitude des Zonalwindes an beiden Stationen mit der Tiefe des winterlichen Polarwirbels (Jacobi, 2000). Damit liegt

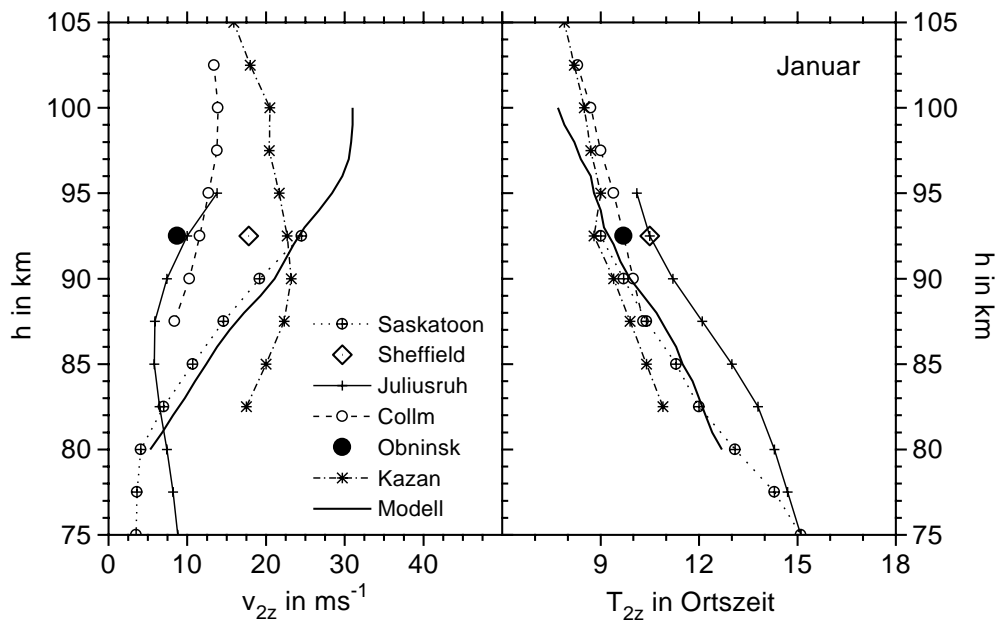


Abbildung 1: Profile der Langzeitmittel der halbtägigen Gezeitenamplitude (links) und Phase (rechts) für Januar an den verschiedenen Messstationen in Tab. 1 zwischen 52°N und 56°N (nach Jacobi et al., 1999).

Meßstation	Koordinaten
Saskatoon	52°N, 107°W
Sheffield	53°N, 4°W
Juliusruh	55°N, 13°E
Collm	52°N, 15°E
Obninsk	55°N, 37°E
Kazan	56°N, 49°E

Tabelle 1: Koordinaten der Meßstationen im Breitenbereich 52 – 56°N.

die Vermutung nahe, daß die beobachteten längenabhängigen Variationen im wesentlichen auf den Einfluß der stationären Wellen, verbunden mit dem winterlichen Polarwirbel in der Stratosphäre, zurückzuführen sind.

2. Stationäre Wellenanregung

Zur Untersuchung des Einflusses des winterlichen Polarwirbels auf das Windfeld in der oberen Mesosphäre und unteren Thermosphäre werden Modellversuche mit einer stationären Wellenanregung durchgeführt. Realisiert wird dies im Modell durch Vorgabe einer Geopotentialstörung am unteren Rand, die zur Vermeidung von numerischen Instabilitäten, gleich einem allmählich wachsenden Gebirge, nach einer Einschwingphase von 40 Tagen eingeschaltet wird.

Die Gleichung für die Störung lautet:

$$\Phi = \tilde{\Phi}(\lambda, \phi) [1 - \exp(-t/\tau)],$$

mit
$$\tilde{\Phi}(\lambda, \phi) = f1(\lambda) \cdot g1(\phi) + f2(\lambda) \cdot g2(\phi),$$

wobei die Gesamtstörung durch ihre Geopotentialanteile zur zonalen Wellenzahl 1 und 2 approximiert wird und f und g die Längen- und Breitenabhängigkeit der Welle 1 und Welle 2-Anteile angeben. Die Längenabhängigkeit wird durch Kosinusfunktionen mit einer zusätzlichen Phase dargestellt, die Breitenabhängigkeit wird im Bereich $0 < \phi \leq 77.5^\circ N$ durch eine Gaussfunktion approximiert, nördlich davon zum Pol hin durch einen linearen Abfall auf Null:

$$\begin{aligned} f1(\lambda) &= \hat{\Phi}_1 \cos(\lambda + \delta\lambda_1), \\ f2(\lambda) &= \hat{\Phi}_2 \cos(\lambda + \delta\lambda_2), \end{aligned}$$

$$g1(\phi) = \begin{cases} \exp(-2((\phi - \delta\phi_1)/\alpha)^2) & : 0 < \phi \leq 77.5^\circ N \\ (90 - \phi)/(90 - 77.5) \cdot g1(77.5^\circ N) & : 77.5 < \phi < 90^\circ N, \end{cases}$$

$$g2(\phi) = \begin{cases} \exp(-2((\phi - \delta\phi_2)/\beta)^2) & : 0 < \phi \leq 77.5^\circ N \\ (90 - \phi)/(90 - 77.5) \cdot g2(77.5^\circ N) & : 77.5 < \phi < 90^\circ N, \end{cases}$$

λ, ϕ : geographische Länge, geographische Breite,
 $\hat{\Phi}_1$: Amplitude der stationären Welle 1 am unteren Rand,
 $\hat{\Phi}_2$: Amplitude der stationären Welle 2 am unteren Rand,
 $\delta\lambda_1$: Phasenlage der stationären Welle 1 am unteren Rand,
 $\delta\lambda_2$: Phasenlage der stationären Welle 2 am unteren Rand,
 $\delta\phi_1$: Phasenlage der stationären Welle 1 am unteren Rand,
 $\delta\phi_2$: Phasenlage der stationären Welle 2 am unteren Rand,
 α, β : Parameter zur Anpassung der Gauss-Funktion.

Die Zeitkonstante τ beträgt einen Tag. Die Modellläufe umfassen einen Zeitraum von weiteren 40 Tagen, so daß sich die Störung am 80. Modelltag voll im Modellgebiet entwickelt hat. Die Phasenlagen und Amplituden der Welle 1 und Welle 2-Anteile der Geopotentialstörung sowie die Parameter der Gaussfunktion sind so gewählt, daß die stationäre Welle in 30 hPa, die Ergebnisse aus Langzeitbeobachtungen möglichst gut approximiert. Daraus ergibt sich:

$$\begin{aligned}
 \hat{\Phi}_1 &= 300 \text{ gpm} \\
 \hat{\Phi}_2 &= 260 \text{ gpm} \\
 \delta\lambda_1 &= +22^\circ E; \quad \delta\phi_1 = 55^\circ N \\
 \delta\lambda_2 &= +125^\circ E; \quad \delta\phi_2 = 60^\circ N \\
 \alpha &= 28.8 \\
 \beta &= 25.5
 \end{aligned}$$

Abb. 2 zeigt die Amplitude und Phase der stationären Geopotentialstörungen zur zonalen Wellenzahl 1 und 2 der Berliner 30 hPa Analysen für Januar in den Jahren 1979 - 1999, sowie die Geopotentialstörung im nächstgelegenen Höhenniveau im COMMA-Modell in 25.8 km, das ca. 25 hPa entspricht. Die Meridionalstruktur der Welle 1 und Welle 2 im Modell stimmt gut mit den Meßwerten überein.

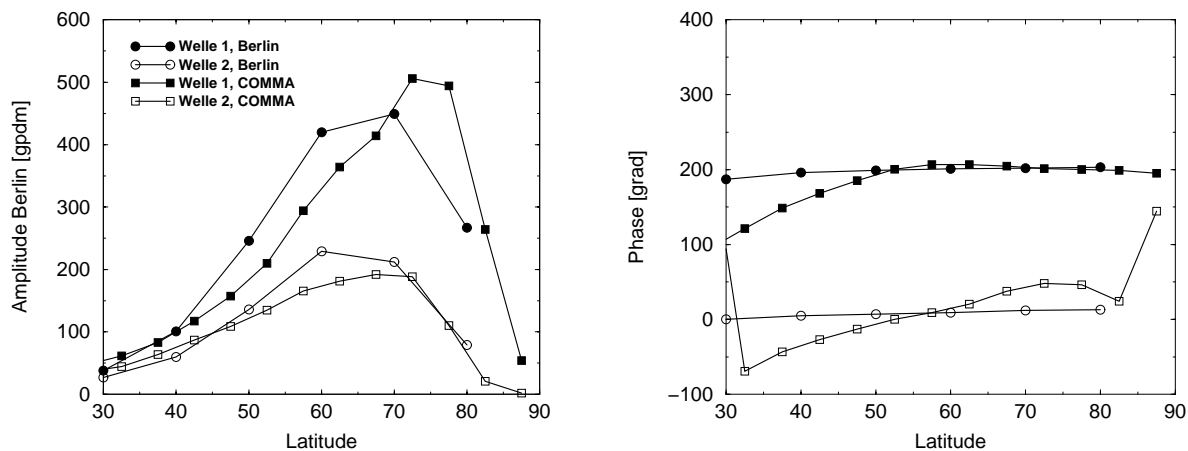


Abbildung 2: Amplitude und Phase der stationären Welle 1 und Welle 2 bestimmt aus 30 hPa Druckflächenhöhen von Berliner Analysen für das Januarmittel im Zeitraum 1979-1999 und stationäre Wellenanteile im Modellauf.

Die Phase der Modellanregung ist so gewählt, daß sie in $50 - 60^\circ N$ mit den Messungen übereinstimmt. Zu niederen Breiten hin wandert die Phase im Modell retrograd, wobei der Einfluß aufgrund der abnehmenden Amplitude jedoch gering ist. Auch die Abweichung der Phase der Welle 2 am Pol ist dort wegen der geringen Amplitude vernachlässigbar. Damit werden in den mittleren Breiten möglichst realistische Bedingungen für die Modellsimulation geschaffen.

3. Modellergebnisse

Abb. 3 zeigt die Modellergebnisse für die stationäre Geopotentialstörung der zonalen Wellenzahl 1 und 2 vor dem Hintergrund des Zonalwindes. Das zonale Windfeld bei Winterbedingungen in der Nordhemisphäre ist charakterisiert durch den stratosphärischen / mesosphärischen Westwindjet in der Winterhemisphäre und dem Ostwindjet in der südlichen Sommerhemisphäre mit den darüberliegenden entgegengerichteten Jets in der unteren Thermosphäre. Es wird eine Zunahme der Geopotentialstörungen im Bereich des Mesosphärenjets beobachtet. Der Welle 1-Anteil erreicht dabei, trotz vergleichbarer Amplitudenstärke in der Anregung, deutlich höhere Werte als der Welle 2-Anteil. Oberhalb des Maximums des Mesosphärenjets wird die stationäre Welle stark gedämpft. Im Bereich der Windumkehr in ca. 90 km Höhe hat die Welle ein Minimum. Oberhalb der Mesopause nimmt die Amplitude der Welle wieder zu. Sie breitet sich auch in den Bereich des Westwindjets in die Südhemisphäre aus und verstärkt sich dort. Die Nullwindlinie stellt also keine strenge kritische Linie für stationäre Wellen im klassischen Sinne dar, die die Wellenausbreitung vollständig blockiert.

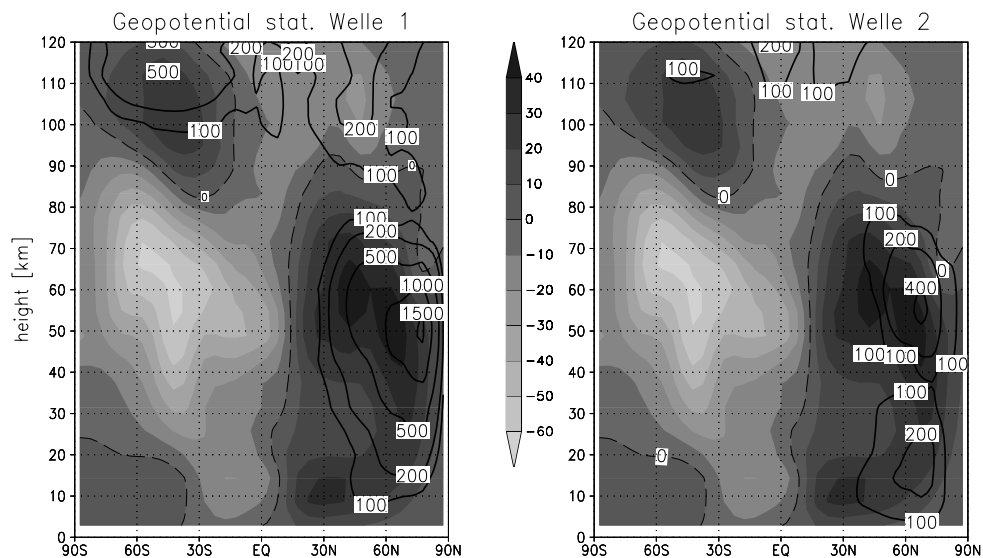


Abbildung 3: Geopotential der stationären Welle 1 und Welle 2 als Konturplot in gpdm vor dem mittleren Zonalwind bei Januarbedingungen. Die Windgeschwindigkeiten sind schattiert dargestellt mit der gestrichelten Nullwindlinie. Das Konturintervall für den Zonalwind ist im Mittelbalken in m/s angegeben, mit positiven Werten für Westwind.

Das Ausbreitungsverhalten der Welle von der mesosphärischen Westwindzone in der Nordhemisphäre in die Westwindbereiche der Südhemisphäre in der unteren Thermosphäre ist bislang nicht geklärt. Das häufig beobachtete Verhalten für Rossby-Wellen, die sich entlang der Nullwindlinie, die als Wellenleiter fungiert, über den Äquator hinweg in die andere Hemisphäre ausbreiten ist hier wegen den trennenden Ostwinden nicht möglich. Mögliche Ursachen können in den Abweichungen von den klassischen Bedingungen liegen, wie beispielsweise dem Einfluß der Einschwingphase des Systems auf die Störung, oder Dämpfungsprozesse, die im Modell enthalten sind. Auch der Bereich der nichtlinearen Wechselwirkungen mit solaren Gezeiten, oder die Modulation von Schwerewellen, sind

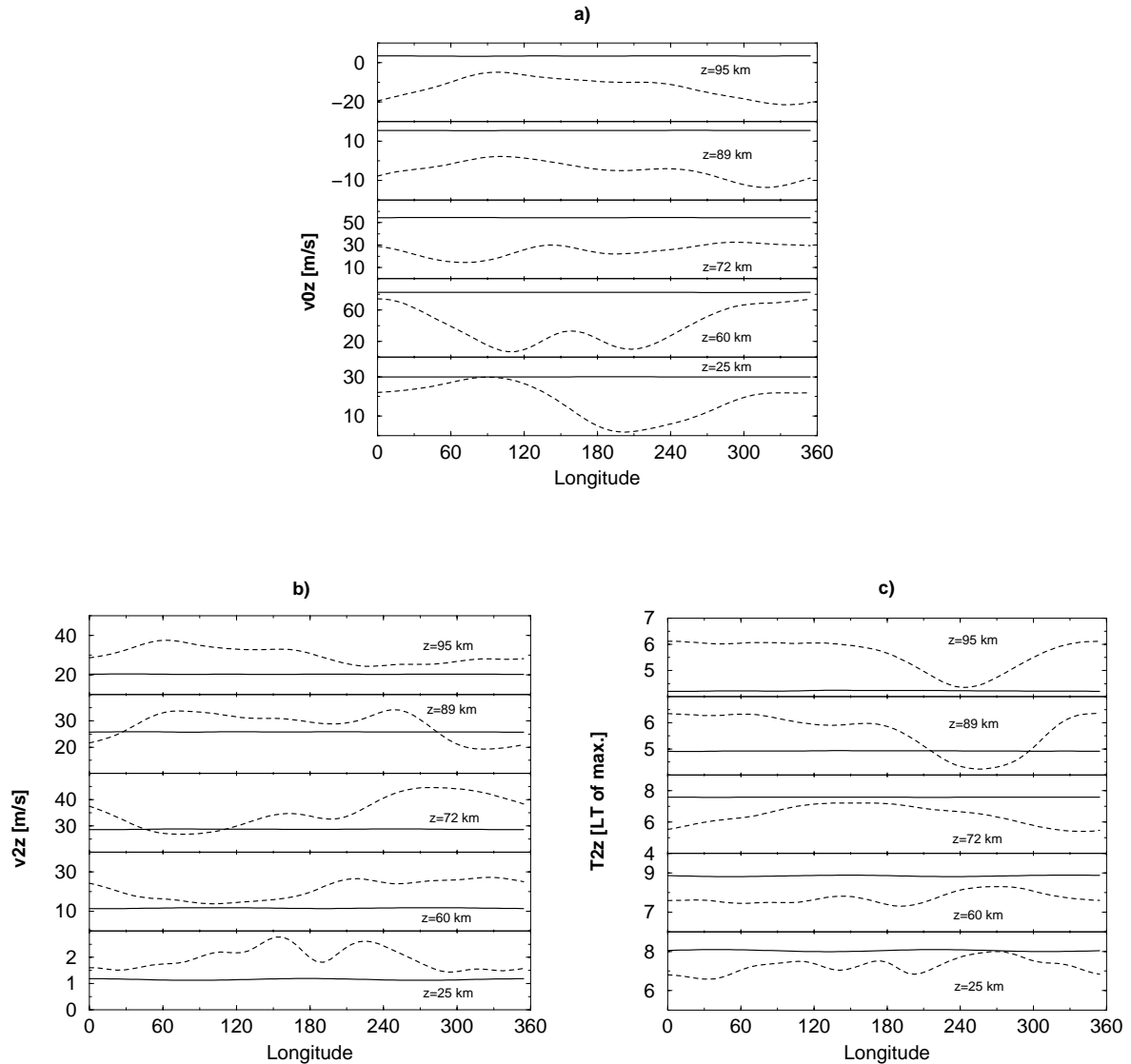


Abbildung 4: Längenabhängigkeit des mittleren Zonalwindes, sowie der Amplitude (b) und Phase (c) des halbtägigen Gezeitenwindes in $52.5^\circ N$ im Kontrolllauf (dicke durchgezogene Linie) und mit Anregung der stationären Welle (gestrichelt). Die Phase ist in der Lokalzeit des Maximums angegeben.

hier noch weiter zu untersuchen. Die Ausbreitung der Welle in die mesosphärischen Ostwinde der Südhemisphäre ist dagegen vollständig blockiert.

Den Einfluß der stationären Welle auf die Längenabhängigkeit des Zonalwindes und der halbtägigen Gezeiten des Zonalwindes in $52.5^\circ N$ zeigt Abb. 4. Der Zonalwind (a) zeigt durch die Wellenstörung im Gegensatz zu dem Kontrolllauf, eine sehr ausgeprägte Längenabhängigkeit mit Variationen von über 20 m/s in 25 km Höhe, knapp 80 m/s im Bereich des Mesosphärenjets in 60 km Höhe und $10 - 20\text{ m/s}$ im Mesopausenbereich. Auch der Mittelwert der Zonalgeschwindigkeit ändert sich in allen Höhen, mit einer starken Abnahme in der Mesosphäre und einer Verstärkung oberhalb der Windumkehr in der Mesopause.

Die halbtägige Gezeitenamplitude (b) zeigt im Modellauf mit Wellenstörung gegenüber dem Kontrolllauf, der wie bei der Zonalgeschwindigkeit eine rein zonalsymmetrische Verteilung aufweist, in allen Höhen oberhalb der Stratosphäre längenabhängige Variationen zwischen 10 und 20 m/s . Darüberhinaus wird in nahezu allen Höhen eine deutliche Verstärkung des Zonalmittels der Amplitude durch die Wellenstörung beobachtet. Auch die Phase (c) zeigt starke längenabhängige Variationen. Diese liegen in der Mesosphäre bei $1 - 1.5$ Stunden mit einer Verschiebung des Mittelwertes um ca. 1 Stunde zu früheren Zeiten. In der Mesopause liegen die Phasenvariationen bei $1.5 - 2$ Stunden, wobei der zonale Mittelwert hier gegenüber dem Kontrolllauf um eine Stunde nachholt.

4. Vergleich mit Messungen

Abb. 5 zeigt einen Vergleich der Modellergebnisse für den halbtägigen zonalen Gezeitenwind im Höhenbereich $89 - 95\text{ km}$ mit Langzeitbeobachtungen im Mesopausenbereich zwischen 52 und $56^\circ N$ in 92.5 km Höhe (Vgl. auch Abb. 1).

Die Absolutbeträge der Phase unterscheiden sich zwischen Modell und Messungen zwar deutlich, die Änderung durch die stationäre Welle in 92.5 km Höhe zeigt jedoch in beiden Fällen eine positive Phasenverschiebung um ca. 1.5 Stunden von Saskatoon zum Nullmeridian. Der Abfall der Phase, der weiter östlich liegenden Stationen wird in den Modellergebnissen allerdings nicht wiedergegeben. Die Variation der halbtägigen Gezeitenamplitude von $10 - 20\text{ m/s}$ in 89 km liegt in der Größenordnung der Messungen. Auch wird die zonale Verteilung mit der ca. 10 m/s stärkeren Amplitude in Saskatoon (Kanada) gegenüber der Station Sheffield richtig wiedergegeben. Die weitere Abnahme zu den östlich liegenden Stationen wird jedoch nicht erfaßt, wohingegen der Anstieg zur östlichsten Station Kazan bei $49^\circ E$ auch in der Simulation beobachtet wird.

Es zeigt sich, daß der langwellige Anteil der Variation in der Simulation relativ gut wiedergegeben wird, wohingegen die Variationen zwischen nahe beieinander liegenden Stationen nicht erfaßt werden. Dies ist bei den großskaligen Störungen der Welle 1 und Welle 2 auch zu erwarten. Die Variationen zwischen Stationen mit einem kurzen Abstand sind möglicherweise auf kleinskaligere Störungen wie z.B. auf die längenabhängige Schwerewellenaktivität zurückzuführen. Berücksichtigt man die Meßungenauigkeit der verschiedenen Stationen zwischen 0 und $40^\circ E$, die mit $5 - 7\text{ m/s}$ angegeben werden (Shepherd, 1996), so liegt die Modellvariation jedoch weitgehend im Bereich der Meßgenauigkeit. Auch bei Ver-

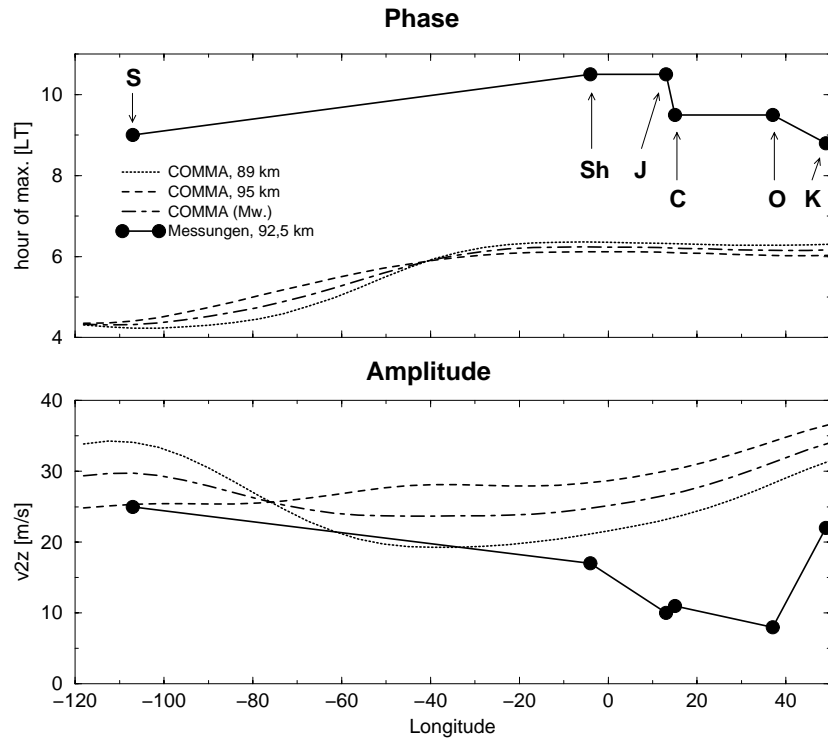


Abbildung 5: Phase und Amplitude der halbtägigen Gezeiten des Zonalwindes in 92.5 km an den verschiedenen Meßstationen, sowie die Modellergebnisse für $z = 89$ km und 95 km und dem Mittelwert in $52.5^\circ N$ bei stationärer Wellenanregung. Die Abkürzungen geben die verschiedenen Meßstationen an. S: Saskatoon ($52^\circ N$, $107^\circ W$), Sh: Sheffield ($53^\circ N$, $4^\circ W$), J: Juliusruh ($55^\circ N$, $13^\circ E$), C: Collm ($52^\circ N$, $15^\circ E$), O: Obninsk ($55^\circ N$, $37^\circ E$), K: Kazan ($56^\circ N$, $49^\circ E$).

gleichem der solaren Gezeiten zwischen Radarwindmessungen und dem Global Scale Wave Model (GSWM), das derzeit als "state of the art" bei der Simulation der Gezeiten in der oberen Mesosphäre und unteren Thermosphäre gilt, gelten Abweichungen in der Größenordnung noch als "good" (Manson et al., 1999; Hagan et al., 1997, 1999). Abweichungen zwischen dem halbtägigen Gezeitenwind im Mesopausenbereich im COMMA-Modell in $52.5^\circ N$ bei Sommerbedingungen und den Langzeitwindmessungen von Collm wurden jedoch bereits auch in früheren Untersuchungen beobachtet (Lange et al. 1999).

Diskussion und Ausblick

Die Untersuchung des Einflusses von stratosphärischen stationären Wellen auf das Windfeld im Bereich der Mesopause in 90 km Höhe hat gezeigt, daß die Variationen, die in Langzeitmessungen entlang des Breitengürtels zwischen $52^\circ N$ und $56^\circ N$ im Winter beobachtet werden, durch die Simulationen in der Größenordnung richtig reproduziert werden. Darüber hinaus sind bei der halbtägigen Gezeitenamplitude auch quantitative Übereinstimmungen in der Phasenlage der Maxima und Minima entlang des Breitenkreises vorhanden. Zwar läßt sich aus den relativ wenigen Beobachtungen noch nicht eindeutig schließen, daß das Ausbreitungsverhalten der stationären Wellenstörung von der Strato-

sphäre in die Mesopausenregion in den Simulation richtig beschrieben wird, die Ergebnisse lassen jedoch den Schluß zu, daß die längenabhängigen Variationen im Winter wesentlich durch den Polarwirbel bedingt sind.

Eine weiterhin offene Frage bleibt, worauf die relativ großen Unterschiede der halbtägigen Gezeitenamplitude zwischen den nahe beieinander liegenden Stationen in Europa zurückzuführen sind. Abgesehen von der Meßungenauigkeit kommen die folgenden Ursachen in Betracht: Zum einen können längenabhängige Variationen der Schwerewellenaktivität durch Wechselwirkungsprozesse starken Einfluß auf die lokalen Eigenschaften der Gezeiten nehmen. Ein weiterer Punkt ist der Einfluß von längenabhängigen Ozonvariationen, die zu Variationen im Antrieb der solaren Gezeiten führen. Weiterhin ist das Ausbreitungsverhalten der stationären Wellen im kritischen Übergangsbereich zwischen Westwind- und Ostwindzonen in der Mesosphäre und unteren Thermosphäre noch unklar, was für das Verständnis der Störung im Mesopausenbereich von Bedeutung ist. Hierzu sind weitere Modellstudien zum Ausbreitungsverhalten und dem Energietransport der Welle erforderlich. Aber auch weitere koordinierte Meßkampagnen zwischen den verschiedenen Stationen tragen zu einer höheren Verlässlichkeit der bisherigen Erkenntnisse bei.

Dank

Diese Untersuchung wurde von der Deutschen Forschungsgesellschaft (Aktenzeichen JA 836/3-2) unterstützt.

Literatur

- Bremer, J., R. Schminder, K.M. Greisiger, P. Hoffmann, D. Kürschner, and W. Singer, 1997: Solar cycle dependence and long-term trends in the wind field of the mesosphere and thermosphere. *J. Atmos. Solar-Terr. Phys.* **59**, 497-509.
- Hagan, M. E., J. L., Chang, and S. K. Avery, 1997: GSWM estimates of non-migrating tidal effects. *J. Geophys. Res.* **102**, 16,439-16,452.
- Hagan, M. E., M. D. Burrage, J. M. Forbes, J. Hackney, W. J. Randel, and X. Zhang, 1999: GSWM-98: Results for migrating solar tides. *J. Geophys. Res.* **104**, 6813-6828.
- Jacobi, Ch., Yu.I. Portnyagin, T.V. Solovjova, P. Hoffmann, W. Singer, A.N. Fahrutdinova, R.A. Ishmuratov, A.G. Beard, N.J. Mitchell, H.G. Muller, R. Schminder, D. Kürschner, A.H. Manson, and C.E. Meek, 1999: Climatology of the semidiurnal tide at 52 – 56°N from ground-based radar wind measurements 1985-1995. *J. Atmos. Solar-Terr. Phys.* **61**, 975-991.
- Jacobi, Ch., 2000: Midlatitude Mesopause Region Dynamics and its Coupling with Lower and Middle Atmospheric Processes, *Wissenschaftliche Mitteilungen aus dem Institut für Meteorologie der Universität Leipzig und dem Institut für Troposphärenforschung e. V. Leipzig*, Bd. **20**, 132 pp.
- Lange, M., R. Schminder, Ch. Jacobi, F., Baier, and G. Günther, 1999: Simulation of middle atmosphere winds and comparison with mesopause region wind measurements. *Adv. Space Res.* **24**, 1527-1530.
- Manson, A., Ch. Meek, M. Hagan, Ch. Hall, W. Hocking, J. MacDougall, S. Franke, D. Riggins, D. Fritts, R. Vincent, and M. Burrage, 1999: Seasonal variations of the

semi-diurnal and diurnal tides in the MLT: multi-year MF radar observations from 2 to 70°N, and the GSWM tidal model. *J. Atmos. Terr. Phys.* **61**, 809-828.

Shepherd, G. 1996: Planetary scale mesopause observing system. Project Description, 1996, 13 pp.

Anschrift des Verfassers:

Martin Lange, Institut für Meteorologie, Stephanstr. 3, 04103 Leipzig.
Email: mlange@uni-leipzig.de

Das lokale Optimierungsproblem

Harald Heinrich, Werner Metz

Zusammenfassung

Es erfolgt eine Darstellung der theoretischen Grundlagen, die für die Berechnung von optimal wachsenden Störungen nötig sind. Bei Anwendung des Projektionsoperators ist es möglich, die Anfangsstörung zu berechnen, welche die größtmögliche kinetische Energie nach Ablauf eines Optimierungsintervalls in einem vorher spezifizierten Gebiet erlangt. Die Berechnung dieser sogenannten lokalen optimalen Vektoren erfolgt für ein barotropes Modell, das um zeitunabhängige Grundströme linearisiert ist.

Die lokalen optimalen Vektoren sind für Grundströme ermittelt worden, die aus einem gleitenden 10-Tages Mittel bestehen. Als Zielgebiet kam dabei der Nordatlantik-Europa Bereich zur Anwendung. Auf diese Weise war es möglich, zu untersuchen, welche Gebiete einen Einfluss auf das Störungswachstum besitzen. Dabei stellte sich heraus, daß für Optimierungszeiten größer 96 Stunden vier verschiedene Startgebiete existieren. Besonderes Interesse galt dem Transport von kinetischer Energie aus diesen Startgebieten in das Zielgebiet. Teilweise zeigte sich, daß die Energie über PNA ähnliche Strukturen in das Zielgebiet getragen wird.

Summary

Under the use of a projection operator it is possible to calculate perturbations with maximal energy growth during optimization time at some pre-chosen geographical area. The so called local optimal perturbations are computed for a barotropic model, linearized about time independent basic states.

The optimal perturbations are determined for basic states consisting of 10-day running means of the daily streamfunctions. The chosen target area was the North Atlantic-European sector. In this way it was possible to find relevant "starting regions" leading to optimal energy growth in the target area. We found four such regions for long optimization times (> 96 hours). Thereby, the mechanisms responsible for the transport of the perturbations from the starting regions into the target area was thereby of special interest. In particular cases we found that structures similar to Pacific North America (PNA) patterns are involved in the energy transfer.

1 Einleitung

Die atmosphärischen Modelle bestehen im allgemeinen aus nichtlinearen partiellen Differentialgleichungen. Die Lösungen solcher Gleichungen besitzen die Eigenschaft, daß

sie sehr empfindlich von den Anfangsbedingungen abhängen. D.h.; kleine Unterschiede in ihnen, können zu Lösungen führen, die miteinander keine Ähnlichkeiten mehr besitzen Lorenz [7]. Für diese Modelle ist es möglich, die Anfangsstörung zu berechnen, die nach Ablauf eines begrenzten Zeitintervalls (Optimierungsintervall) die größtmögliche Energie besitzt Lorenz [8], Farrell [5], Borges u. Hartmann [2], Molteni u. Palmer [11]. Diese Störungen werden als optimale Vektoren oder auch singuläre Vektoren bezeichnet. Die optimalen Vektoren lassen sich sowohl für stabile als auch für instabile Grundströme berechnen. Diese Tatsache macht sie gerade für die Anregung von atmosphärischen Anomalien interessant.

Zum einen ist es möglich, die Optimierung global, d.h. für die gesamte Atmosphäre durchzuführen. Borges und Hartmann [2] berechneten die globalen optimalen Vektoren für ein barotropes Modell. Ihr Ziel war es, zu klären, ob diese Störungen als Anregung für langperiodische Anomalien fungieren. Ein Vertreter dieser Anomalien sind die blockierenden Hochs. Sie besitzen die Eigenschaft, sich in einem festen geographischen Gebiet auszubilden und über einen längeren Zeitraum zu bestehen. Ihre Entstehungszeiten sind relativ kurz und somit scheinbar auf die optimalen Vektoren zugeschnitten. Durch nichtlineare Integration ihres Modells konnten sie die beobachteten Anomalien sehr gut nachbilden. Als Eingangsdaten dienten dabei die gemittelten Grundströme, denen sie die optimalen Störungen überlagerten. Molteni und Palmer [11] führten ebenfalls eine globale Optimierung durch. Dafür verwendeten sie ein barotropes und ein baroklines 3-Schichten Modell. Ein Vergleich der optimalen Vektoren beider Modelle ergab, daß die baroklinen Störungen für längere Optimierungsintervalle ein größeres Wachstum erfahren, da sie die barotrope und barokline Energieumwandlung nutzen. Unter der Annahme, daß der erste optimale Vektor durch ein Rossbywellenpaket darstellbar ist, beschrieb Oortwijn [12] die barotrope Ausbreitung dieses Paketes mit einer WKB Approximation. Für zonale aber auch nichtzonale Grundströme untersuchte er unter Zuhilfenahme von Erhaltungssätzen, auf welche Art und Weise Energie vom Grundstrom in Störungsenergie der optimalen Vektoren übergeht. Er konnte zeigen, daß das Störungswachstum in zwei Phasen vonstatten geht. In der Anfangsphase dominieren das Störungswachstum barokline Prozesse, nach deren Ablauf die barotrope Phase des Störungswachstums vorherrscht.

Auf der anderen Seite ist es möglich, die Anfangsstörungen zu berechnen, die in einem Teilgebiet der Atmosphäre die größtmögliche Energie erlangen. Barkmeijer [1] löste diese lokale Optimierung für ein barotropes Modell unter Verwendung des adjungierten linearisierten Modells. Dazu führte er einen Projektionsoperator ein, der es gestattet, die kinetische Energie eines Teils der Atmosphäre zu ermitteln. Ziel seiner Arbeit war die Analyse der Stabilitätsverhältnisse von sogenannten Rossby-Haurwitz Wellen (sie sind Lösungen des barotropen Modells). Für realistische Grundströme führte Buizza [3] die lokale Optimierung durch. Er verwendete ein baroklines 19-Schichten Modell und maximierte die gesamte Störungsenergie. Dazu nutzte er die sogenannte totale Energienorm Buizza et al. [4]. Für Ensemble Vorhersagen wollte Buizza die optimalen Vektoren als Anfangsbedingungen wählen. Die Nordatmosphäre kann im Sommer hinreichend stabil sein, so das bei einer globalen Optimierung Störungen, die sich auf

der Südhalbkugel befinden, dominieren. Durch diese Probleme motiviert, berechnete er die lokalen optimalen Vektoren der Nordhemisphäre. Dabei verwendete er für die Berechnung der lokalen Energie ebenfalls einen Projektionsoperator, den er, um die Symmetrie des Problems zu wahren, zweimal anwenden mußte.

Diese Zweifachanwendung des Projektors kann, bei kurzen Optimierungsintervallen, zu Ungenauigkeiten führen. Um dies zu vermeiden, stellt die vorliegende Arbeit einen Weg vor, die lokale Optimierung durch einmaliges Anwenden des Projektionsoperators unter Beibehaltung der Symmetrie zu lösen. Eine Demonstration des Verfahrens erfolgt anhand eines barotropen Modells, welches um zeitlich konstante Grundströme linearisiert ist. Als ein erstes Beispiel soll die Optimierung für die Südhalbkugel dienen. Hieran ist es möglich, die Prozesse die barotropes Störungswachstum nach sich ziehen, zu erkennen. Weiterhin erfolgt die lokale Optimierung für den Nordatlantik-Europa Bereich. Um zu testen, welche Ursprungsgebiete für das Störungswachstum einen Einfluß auf dieses Zielgebiet besitzen, kamen Grundströme, welche eine Mittelung über 10-Tages Intervalle erfuhren, zum Einsatz. Gerade die Mechanismen, die für den Transport von Störungsamplitude in des Optimierungsgebiet verantwortlich sind, waren von besonderem Interesse. Dabei spielt das Vermögen der optimalen Vektoren, langperiodische Anomalien anzuregen, ebenfalls eine große Rolle.

Die Arbeit ist folgendermaßen gegliedert. In Abschnitt 2 werden die verwendeten Daten vorgestellt. Abschnitt 3 geht auf das barotrope Modell und die theoretischen Grundlagen ein. Abschnitt 4 stellt die Ergebnisse vor, für die schließlich in Abschnitt 5 eine Diskussion erfolgen soll.

2 Daten

Für die Untersuchungen im zweiten Teil, standen die Stromfunktionen 300 hPa für sechs Nordwinter zur Verfügung. Jeder Winter besteht aus 110 Tagen der täglichen 00 UTC Analysen. Starttag ist jeweils der 20. November. Die Daten, die die Periode von 1979/80 bis 1985/86 umfassen, stammen vom European Centre for Medium Range Weather Forecast (ECMWF). Sie sind nach Kugelflächenfunktionen entwickelt. Die dabei verwendete Dreiecksabhackung bei einer Gesamtwellenzahl von 21 (T21) entspricht einer horizontalen Auflösung von ca. $5.6^\circ \times 5.6^\circ$ auf einem Gaußgitter.

Für den klimatologischen Grundstrom aus Abschnitt 4.1 erfolgt die Mittelung der täglichen Stromfunktionen über die gesamten sechs Winterjahre. In Abschnitt 4.2 erfolgt die Berechnung der lokalen optimalen Vektoren für sogenannte niederfrequente Grundströme. Sie resultieren aus 10-Tägiger Mittelung der Stromfunktionen. Da es vorher nicht bekannt war, welche Grundströme sich gegenüber anderen auszeichnen, kam ein gleitendes Mittel zum Einsatz. Daraus entstehen pro Winter ca. 90 Grundströme, für die das Optimierungsproblem gelöst wurde.

3 Modell und Methode

3.1 Modell

Das hier verwendete barotrope Modell ist ein globales spektrales Modell. Für die dabei auftretenden Approximationen der Geschwindigkeitsfelder mittels Kugelflächenfunktionen kommt eine T21 Abhackung zum Einsatz. Die divergenzfreie barotrope Vorticitygleichung besitzt folgende Gestalt

$$\frac{\partial \zeta}{\partial t} = -\nabla \cdot \mathbf{v} (\zeta + f) - \alpha \zeta - \gamma \nabla^4 \zeta \quad (1)$$

wobei ζ die relative Vorticity, $\mathbf{v} = (u, v)$ der horizontale Geschwindigkeitsvektor, f der Coriolis-Parameter, α der lineare Reibungs- und γ der Diffusionskoeffizient sind. Als Standardwerte für die 300 hPa Druckfläche kommen $\alpha = 1/7 \text{ Tag}^{-1}$ und $\gamma = 2.239 \cdot 10^{16} \text{ m}^4 \text{ s}^{-1}$ zur Anwendung. Die divergenzfreie Windgeschwindigkeit erhält man aus $\mathbf{v} = \mathbf{k} \times \nabla \psi$, ψ wird als Stromfunktion bezeichnet. Aus ihr läßt sich die Vorticity berechnen, es gilt

$$\zeta = \nabla^2 \psi \quad (2)$$

Nach Linearisierung und Diskretisierung Metz [10] erhält man folgendes 480 dimensionale lineare Differentialgleichungssystem

$$\frac{d\mathbf{\Psi}}{dt} = \mathbf{A}(\bar{\psi}) \mathbf{\Psi} \quad (3)$$

mit dem reellen Vektor $\mathbf{\Psi}$, in dem die Kugelfunktionskoeffizienten der Störungsstromfunktion ψ' (abwechselnd Real- und Imaginärteile) angeordnet sind, und der reellen (480×480) -Matrix \mathbf{A} , welche von dem Grundstrom $\bar{\psi}$ abhängt.

3.2 Theorie

Die barotrope Vorticitygleichung läßt sich im allgemeinen folgendermaßen darstellen

$$\frac{\partial \zeta}{\partial t} = F(\zeta) \quad (4)$$

wobei $F(\zeta)$ eine nichtlineare differenzierbare Funktion ist. Linearisierung dieser Gleichung, d.h. die Vorticity wird in einen Grundzustand und eine Abweichung davon zerlegt $\zeta = \bar{\zeta} + \zeta'$, ergibt eine Gleichung die die Entwicklung einer kleinen Störgröße ζ' beschreibt.

$$\frac{\partial \zeta'}{\partial t} = A \zeta' \quad (5)$$

A stellt hierbei einen linearen Differentialoperator dar. Aus der Linearität von (5) folgt, daß mit ψ und ϕ auch $\alpha\psi + \beta\phi \quad \forall \alpha, \beta \in \mathcal{C}$ Lösungen sind. ζ' kann man als Vektor in einem Hilbertraum \mathcal{H} auffassen. Da im barotropen Fall Lösungen auf einer

Kugelschale gesucht werden, besitzt \mathcal{H} folgende Gestalt: $\mathcal{H} = L^2(\Omega, d^2x) = L^2([0, 2\pi] \times [-1, 1], d\mu d\lambda)$ (Raum der Quadratintegrierbaren Funktionen auf der Sphäre) mit dem zugrundeliegendem Skalarprodukt

$$\{ \psi, \phi \} = \int_0^{2\pi} \int_{-1}^1 \psi^* \phi d\mu d\lambda \quad (6)$$

$\forall \psi, \phi \in \mathcal{H}$ und der damit assoziierten Norm

$$\| \psi \|^2 = \int_0^{2\pi} \int_{-1}^1 \psi^* \psi d\mu d\lambda \quad (7)$$

Der hochgestellte Stern bedeutet konjugiert komplex, λ ist die Länge und μ ist der Sinus der Breite. Die Basis von \mathcal{H} stellen Kugelflächenfunktionen dar. Sie besitzen die folgende Gestalt

$$Y_{m,n}(\mu, \lambda) = \frac{1}{\sqrt{2\pi}} P_{m,n}(\mu) e^{im\lambda} \quad (8)$$

mit der Orthogonalitätsrelation

$$\int_0^{2\pi} \int_{-1}^1 Y_{m,n} Y_{m',n'}^* d\mu d\lambda = \delta_{m,m'} \delta_{n,n'} \quad (9)$$

wobei $\delta_{x,y}$ die Deltadistribution ist. Für jeden beliebigen Vektor $\psi \in \mathcal{H}$ gilt

$$\psi(\mu, \lambda, t) = \sum_{m=-N}^N \sum_{n=|m|}^N \psi_{m,n}(t) Y_{m,n}(\mu, \lambda) \quad (10)$$

Die Störungsvorticity ζ' ist über $\zeta' = \nabla^2 \psi$ mit der Störungsstromfunktion ψ verknüpft. Mit ψ und dem Skalarprodukt (6) läßt sich die kinetische Energie einer Störgröße folgendermaßen berechnen [13]

$$E_{kin} = \{ \psi, -\nabla^2 \psi \} \quad (11)$$

Der Laplaceoperator wird hierbei in Kugelkoordinaten (ohne r-Abhängigkeit) verwendet. Der Operator $-\nabla^2$ ist bezüglich diesem SKP selbstadjungiert (siehe Anhang 1). Das ist sinnvoll, da die kinetische Energie eine reelle Größe sein muß.

Mit Hilfe eines Projektionsoperators \hat{T}_g läßt sich die kinetische Energie eines Teilgebietes $g \subset \Omega$ berechnen. \hat{T}_g ist folgendermaßen definiert

$$\hat{T}_g \psi(\mu, \lambda, t) = \chi(\mu, \lambda) \psi(\mu, \lambda, t) \quad (12)$$

wobei

$$\chi(\mu, \lambda) = \begin{cases} 1 & (\mu, \lambda) \in g \\ 0 & (\mu, \lambda) \notin g \end{cases} \quad (13)$$

Mehrfachanwendung des Projektionsoperators ändert nicht das Resultat, d.h. es gilt: $\hat{T}_g^2 = \hat{T}_g \hat{T}_g = \hat{T}_g$. Für die kinetische Energie folgt nun unter Anwendung von \hat{T}_g

$$E_{kin}^g = - \int \int_g \psi^* \nabla^2 \psi \, d\mu \, d\lambda = - \int_0^{2\pi} \int_{-1}^1 \psi^* \hat{T}_g \nabla^2 \psi \, d\mu \, d\lambda = - \{ \psi, \hat{T}_g \nabla^2 \psi \} \quad (14)$$

Wie man hieraus erkennen kann, ist der Projektionsoperator \hat{T}_g ebenfalls selbstadjungiert.

Um nun die kinetische Energie E_{kin}^g eines Gebietes $g \subset \Omega$ zu maximieren, ist es nützlich, die Gleichung (14) unter Einführung der Energiematrix \mathbf{D} umzuformen. Aus den Gleichungen (9), (10) und der Tatsache, daß Kugelflächenfunktionen Eigenfunktionen des Laplaceoperators sind, findet man

$$E_{kin} = \sum_{m=-N}^N \sum_{n=|m|}^N n(n+1) |\psi_{m,n}|^2 = \langle \Psi, \mathbf{D}\Psi \rangle \quad (15)$$

Hierbei bedeutet $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^k x_i y_i$ das reelle Euklidische Skalarprodukt ($(\cdot)^T$ steht für transponieren), \mathbf{D} ist eine reelle Diagonalmatrix mit den Eigenwerten des Laplaceoperators auf der Diagonalen, und Ψ ist ein reeller Vektor, dessen Komponenten die Koeffizienten $\psi_{m,n}(t)$ der Entwicklung (10) enthält. Eine genauere Beschreibung dieser Prozedur befindet sich bei Heinrich [6]. Es zeigt sich später, daß es von Vorteil ist, die Energiematrix folgendermaßen darzustellen: $\mathbf{D} = \mathbf{D}^{1/2} \mathbf{D}^{1/2}$, wobei $\mathbf{D}^{1/2}$ als Diagonalmatrix, die die Wurzel der Diagonalelemente von \mathbf{D} enthält, definiert ist. Unter Ausnutzung der Selbstadjungiertheit von $\mathbf{D}^{1/2}$ und Anwendung des Projektionsoperators findet man für die kinetische Energie eines Teilgebietes g

$$E_{kin}^g = \langle \mathbf{D}^{1/2} \Psi, \mathbf{T}_g \mathbf{D}^{1/2} \Psi \rangle \quad (16)$$

Da man es hierbei mit dem Euklidischen Skalarprodukt zu tun hat, folgt für den Projektionsoperator eine etwas kompliziertere Gestalt. Es ist aber möglich, ihn durch eine symmetrische Matrix \mathbf{T}_g zu repräsentieren (siehe Anhang 2).

3.3 Optimierung für ein lokales Gebiet

Gesucht ist die Anfangsstörung $\Psi(t_0) = \Psi_0$, die nach Ablauf einer zeitlichen Entwicklung (im voraus festgelegtes Optimierungsintervall T_{opt}) die größtmögliche kinetische Energie in einem vorher definierten Gebiet $g \subset \Omega$ erlangt. D.h., es gilt

$$E_{kin}^g(t_0 + T_{opt}) = Max! \quad (17)$$

für $\Psi_0 = \Psi_{opt}$ dem ersten lokalen optimalen Vektor unter der Nebenbedingung

$$E_{kin}(t_0) = \langle \Psi_0, \mathbf{D}\Psi_0 \rangle = 1 \quad (18)$$

Es wird hier $E_{kin}(t_0)$ und nicht $E_{kin}^g(t_0)$ verwendet, da die Anfangsstörung Ψ_{opt} nicht ausschließlich in $g \subset \Omega$ liegen muß. Es ist vielmehr ja denkbar, daß sich die Störung im Zuge ihrer zeitlichen Entwicklung in das Zielgebiet hinein bewegt.

Die Lösung von Gleichung (3) läßt sich folgendermaßen darstellen

$$\Psi(t) = \mathbf{P}(t, t_0) \Psi(t_0) \quad (19)$$

mit dem Propagator $\mathbf{P} = \mathbf{P}(t, t_0)$ und der Anfangsbedingung $\Psi(t_0) = \Psi_0$. Die Frage nach dem Auffinden des optimalen Vektors Ψ_{opt} ist folgendem Eigenwertproblem äquivalent

$$[\mathbf{P}^T \mathbf{D}^{1/2} \mathbf{T}_g \mathbf{D}^{1/2} \mathbf{P} - \lambda_i \mathbf{D}] \Psi_{0_i} = 0 \quad (20)$$

Die Eigenwerte λ_i sind die optimalen Werte, sie lassen sich ihrer Größe nach ordnen; d.h.:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \quad (21)$$

Sie repräsentieren das Verhältnis aus kinetischer Störungsenergie nach Ablauf des Optimierungsintervalles $t = t_0 + T_{opt}$ zu kinetischer Störungsenergie zum Zeitpunkt $t = t_0$. Der erste optimale Vektor Ψ_{opt} ist der zum größten Eigenwert λ_1 gehörige Eigenvektor. Er besitzt nach Ablauf des Optimierungsintervalls die größte kinetische Energie im Gebiet $g \subset \Omega$.

Für die Lösung der Eigenwertgleichung (20) kommt ein Lanczos Algorithmus zum Einsatz Buizza et al. [4]. Diese Methode besitzt den Vorteil, daß die einzelnen Matrixelemente nicht explizit bekannt sein müssen. Es genügt das Wissen über die Wirkung der Matrix auf beliebige Vektoren.

4 Ergebnisse

4.1 Optimale Vektoren der Südhalbkugel

Als Optimierungsgebiet soll als erstes einmal die gesamte Südhalbkugel dienen. Die lokale Optimierung ist notwendig, da im Nordwinter Störungen, die sich ausschließlich auf der Nordhalbkugel befinden, dominieren. Es kommt der in Abschnitt 2 beschriebene klimatologische Grundstrom zur Anwendung. Er besteht aus einem 6-jährigen Mittel der 300 hPa Winterstromfunktionen (Dezember, Januar, Februar). Die Strömung der Südhalbkugel ist in Abbildung 1a in polarer stereographischer Projektion dargestellt. Wie man sehen kann, besitzt sie einen sehr stark zonalen Charakter. Bei genauerem Betrachten ist erkennbar, daß die Isolinien im Bereich 30 Grad westliche bis 150 Grad östliche Länge stärker gedrängt sind als im übrigen Gebiet.

Beispielhaft für die zeitliche Entwicklung ist in Abbildung 2 der erste optimale Vektor für das Optimierungsintervall 48 Stunden dargestellt. Seine Ausgangslage zum Zeitpunkt $t = t_0$ (Abbildung 2a) befindet sich nördlich des Grundstrombereichs, in dem auch die Isoliniendrängung am größten ist. Im Laufe seiner zeitlichen Entwicklung (Abbildungen 2b-d) bewegt er sich nach Osten, wobei er sich gleichzeitig nach Süden

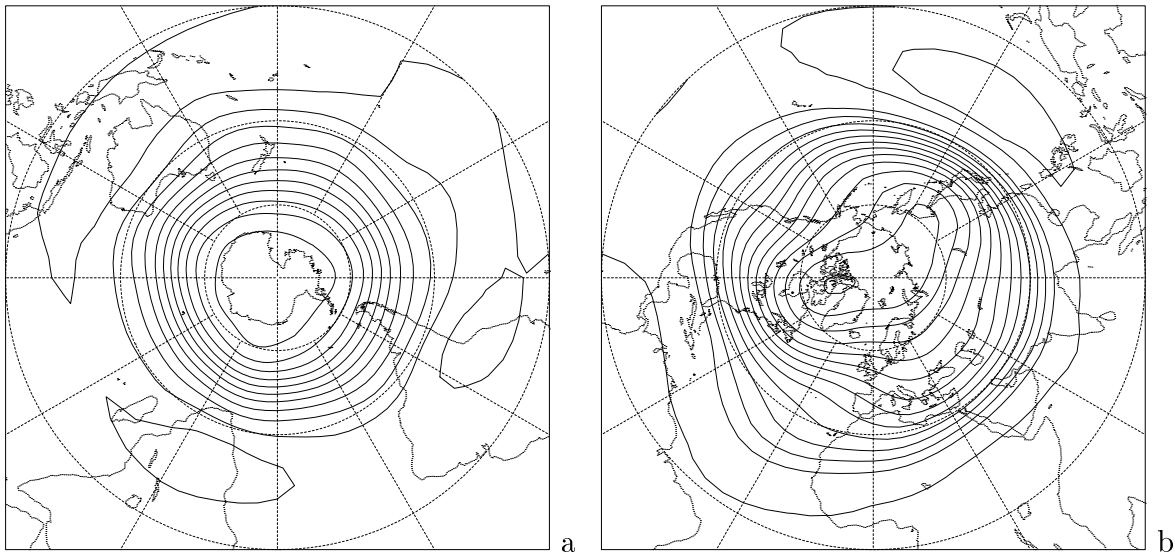


Abbildung 1: Klimatologischer Grundstrom (D,J,F) 300 hPa Stromfunktion in polarer stereographischer Projektion der (a) Südhalbkugel und (b) Nordhalbkugel. Das Zeichnungsintervall beträgt $10 \times 10^6 \text{ m}^2 \text{ s}^{-1}$, positive sowie negative Werte sind mit durchgezogenen Linien dargestellt.

verlagert. Nach 48 Stunden (Abbildung 2c) befindet sich der größte Teil seiner Amplitude in dem Gebiet, in dem auch der Stromfunktionsgradient des Grundstromes am stärksten ist.

Sehr schön sind an diesem Beispiel die Mechanismen zu erkennen, die das Störungswachstum verursachen. Zum einen muß die Störung, um maximal kinetische Energie dem Grundstrom entnehmen zu können, durch das Gebiet größter kinetischer Energie laufen. Zum anderen, kann man sehr schön feststellen, daß die Neigung der einzelnen Strukturen mit der Entwicklungszeit zunimmt. Zum Zeitpunkt $t = t_0$ liegt die Störung fast parallel zur Grundstromachse, nach 48 Stunden stehen sie senkrecht und nach 96 Stunden ist ein Durchknicken einzelner Strukturen zu beobachten. Durch Änderung der Neigung ist es für die Störungen möglich, kinetische Energie zu gewinnen (U. Harlander, persönliches Gespräch). Das Durchknicken in die entgegengesetzte Richtung nach 96 Stunden ist ein Anzeichen dafür, daß die Störungen Energie verlieren; d.h., es geht Störungsenergie in Grundstromenergie über. Die beiden oben beschriebenen Mechanismen führen zu Amplitudenwachstum und einer damit verbundenen Verstärkung der Störungsstromfunktionsgradienten, die proportional zur kinetischen Energie sind.

Wie aus einer EOF-Analyse (Empirische Orthogonal Funktionen) der Südhemisphäre hervorgeht, besteht eine beachtenswerte Ähnlichkeit zwischen dem optimalen Vektor zum Optimierungszeitpunkt (Abbildung 2c) und der zweiten EOF. Dies trifft auf die Lage und die Skala gleichermaßen zu. Beide Strukturen besitzen ca. die zonale Wellenzahl 6, wobei die Wellenberge und Täler fast in Phase liegen. D.h.; in der

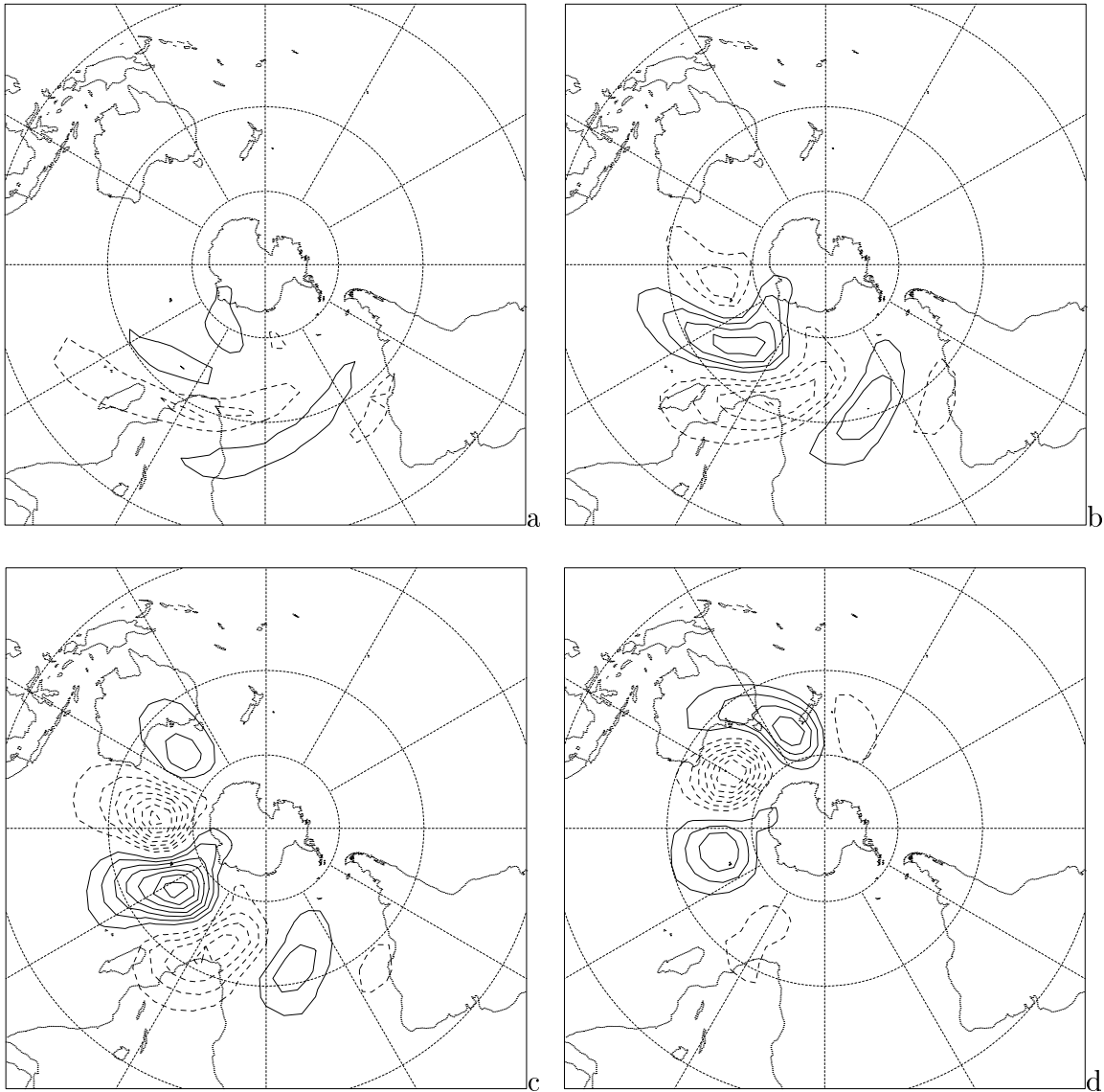


Abbildung 2: erster optimaler Vektor der Südhalbkugel für 48 Stunden (a) zum Zeitpunkt $t = t_0$, (b) nach 24 Stunden, (c) nach 48 Stunden und (d) nach 96 Stunden Entwicklung. Das Zeichnungsintervall beträgt $3 \times 10^5 \text{ m}^2 \text{ s}^{-1}$.

Umgebung der Wellenpaketzentren besteht eine größere Übereinstimmung als an den Rändern. Der Hauptunterschied zwischen dem optimalen Vektor und der zweiten EOF besteht in der genau entgegengesetzten Neigung der Strukturen. Die EOF's, die mit den wandernden synoptikskaligen Störungen (Hoch- und Tiefdruckgebiete der mittleren Breiten) in Verbindung gebracht werden, speisen Energie in den Grundstrom. Entgegengesetzt dazu, nehmen die optimalen Vektoren kinetische Energie des Grundstroms auf. Es wäre denkbar, daß die physikalischen Prozesse die für die Ausbildung

der EOF Muster verantwortlich sind, ebenfalls bei den optimalen Vektoren eine Rolle spielen. Anders herum, erklären die optimalen Vektoren einen großen Teil der Varianz des zugrundeliegenden Grundstromes und könnten damit die Lage der EOF Muster vorgeben.

4.2 Der Atlantik-Europa Bereich

Nachfolgend soll der Einfluß auf das Störungswachstum im Nordatlantik-Europa Bereich aus anderen Gebieten untersucht werden. Dazu erfolgt die Berechnung der lokalen optimalen Vektoren für dieses Gebiet (47 Grad West bis 22 Grad Ost; 34 bis 90 Grad Nord). Als Daten stehen die täglichen Winterstromfunktionen (Dezember, Januar, Februar) in dem Zeitraum 1979 bis 1985 zur Verfügung. Für das Mittelungsschema kommt ein gleitendes 10-Tages Mittel zur Anwendung (siehe Abschnitt 2), welches über die vorhandenen Winterdaten geschoben wird. Pro Winter entstehen somit ca. 90 Grundströme.

Als Optimierungsintervall für die lokalen optimalen Vektoren kommt zunächst einmal $T_{opt} = 48$ Stunden zum Einsatz. Um einen Eindruck über die Start- und Zielgebiete zu erlangen, erfolgt die Transformation der optimalen Vektoren in den Gitterpunktstraum. Hier findet eine Überprüfung der einzelnen Gitterpunktswerte statt. Bei Überschreitung eines vorher definierten Schwellenwertes wird in einem zweiten Gitter, der Häufigkeitsverteilung, an dieser Stelle eine eins addiert. Abbildung 3a zeigt die Häufig-

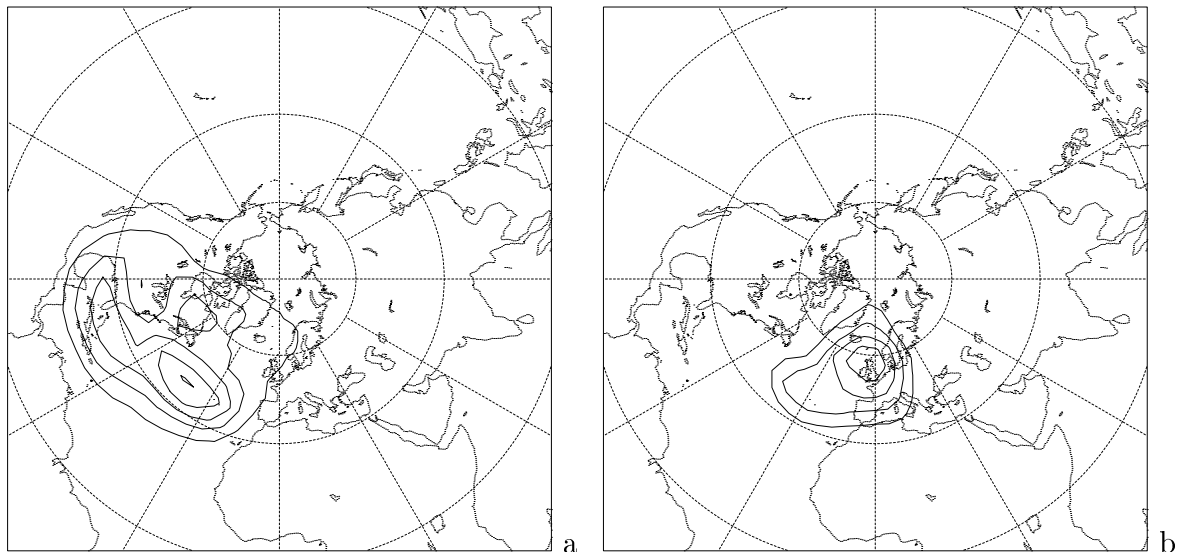


Abbildung 3: Häufigkeitsverteilungen der lokalen ersten optimalen Vektoren für $T_{opt} = 48$ Stunden zum Zeitpunkt (a) $t = t_0$ und (b) $t = T_{opt}$. Das Zeichnungsintervall beträgt 70.

keitsverteilung der optimalen Vektoren zum Zeitpunkt $t = t_0$. Aus einem Vergleich mit dem klimatologischen Grundstrom (Abbildung 1b) geht hervor, daß sie in der Region zu finden sind, in der im Mittel auch der Atlantikjet liegt. Die Verhältnisse nach Ablauf des Optimierungsintervalles sind in Abbildung 3b gezeigt. Wie nicht anders zu erwarten, liegen die optimalen Vektoren jetzt in dem Gebiet, für welches sie optimiert sind.

Vergrößert man das Optimierungsintervall auf $T_{opt} = 192$ Stunden, können vier Startgebiete identifiziert werden (Abbildung 4a). Zwei davon liegen im Atlantikjetbereich und die anderen zwei im Asiatischen Subtropenjetbereich, wobei die Gebiete bezüglich der klimatologischen Jetachsen jeweils nach Süden bzw. nach Norden verschoben sind (vergleiche mit Abbildung 1b). Nach Ablauf der 192 Stunden ist die Häufigkeitsverteilung in Abbildung 4b dargestellt. Sie zeigt große Ähnlichkeit mit der Verteilung für das 48-stündige Optimierungsintervall (Abbildung 3b).

Um die kinetische Störungsenergie im Atlantik-Europa Raum zu maximieren, wird für das Optimierungsintervall 192 Stunden neben dem Atlantikjet auch der Asiatische Subtropenjet genutzt. Die optimalen Störungen verwenden die Jet's zur Energieaufnahme und als Ausbreitungsmedium. Sie müssen dann, während ihrer zeitlichen Entwicklung, diese Energie in das Zielgebiet transportieren. Dieses Verhalten ist einmal in Abbildung 5 dargestellt. Sie zeigt den ersten lokalen optimalen Vektor für das Optimierungsintervall $T_{opt} = 192$ Stunden samt seiner zeitlichen Entwicklung. Der zugrundeliegende Grundstrom besteht aus einem 10-tages Mittel der täglichen Strom-

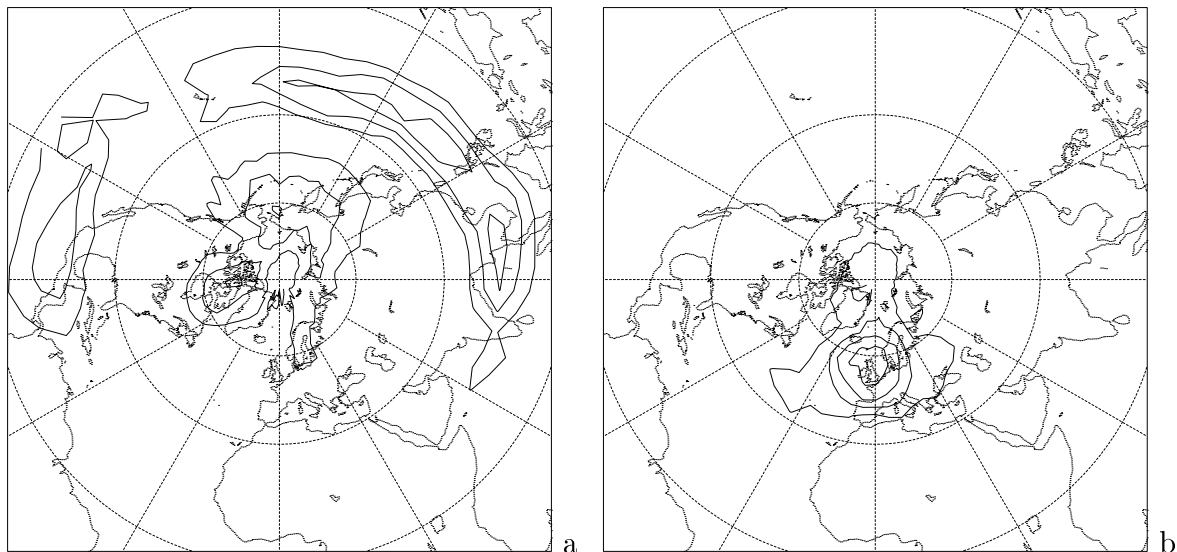


Abbildung 4: Häufigkeitsverteilungen der lokalen ersten optimalen Vektoren für $T_{opt} = 192$ Stunden zum Zeitpunkt (a) $t = t_0$ und (b) $t = T_{opt}$. Das Zeichnungsintervall beträgt 70.

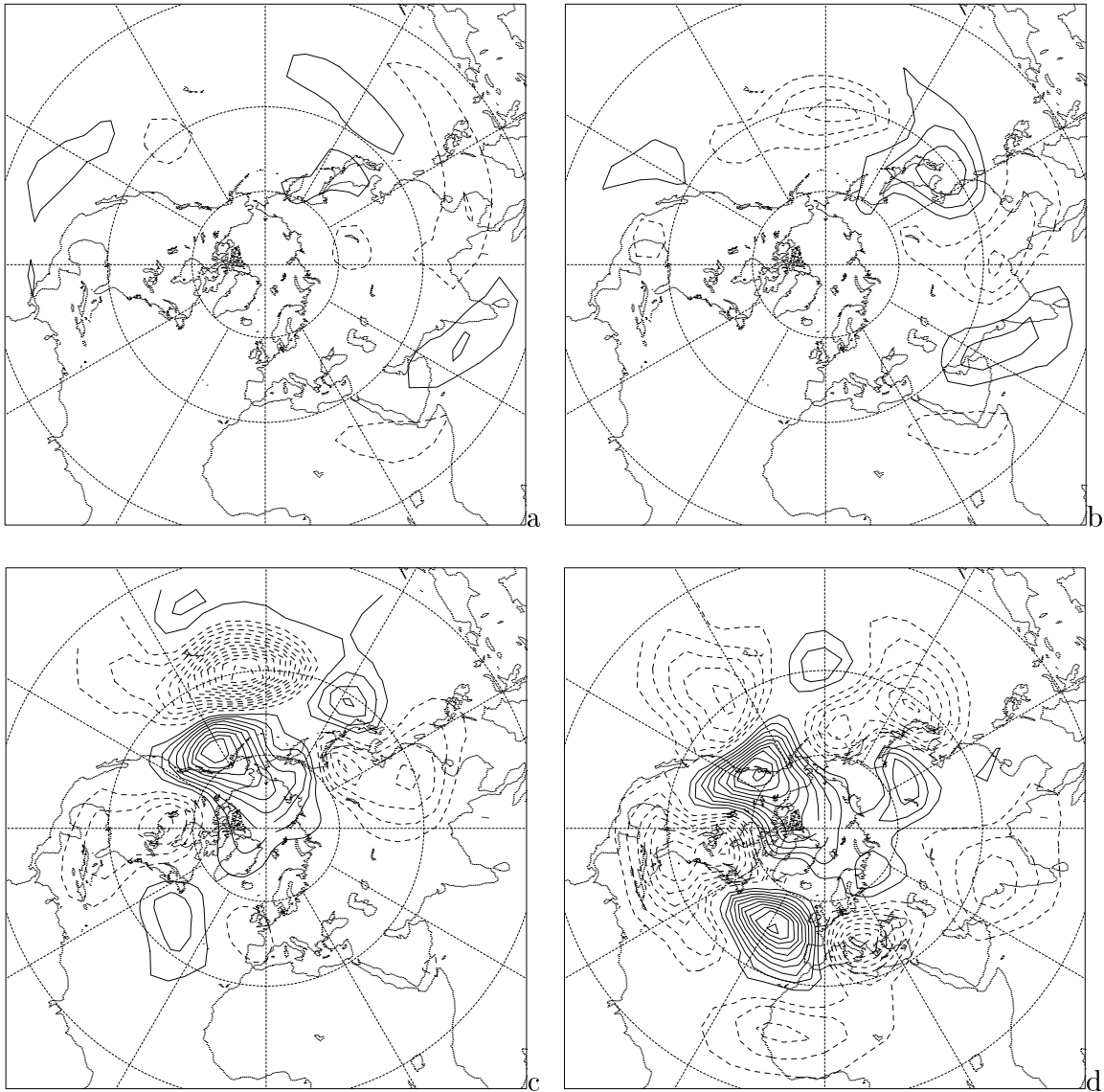


Abbildung 5: Erster lokaler optimaler Vektor des Nordatlantik-Europa Bereichs für den Grundstrom 21.01.84 bis 30.01.84 für 192 Stunden (a) zum Zeitpunkt $t = t_0$, (b) nach 24 Stunden, (c) nach 96 Stunden und (d) nach 192 Stunden Entwicklung. Das Zeichnungsintervall ist in allen vier Grafiken unterschiedlich.

funktionen. Starttag für die Mittelungsperiode ist der 21.01.84. In diesem speziellen Beispiel kann man erkennen, daß der Hauptteil der Amplituden in der Mündungsregion des Asiatischen Subtropenjets liegt (Abbildung 5a). Während ihrer zeitlichen Entwicklung, läuft die Störung durch den Jet und gewinnt dabei kinetische Energie. Nach 96 Stunden Integration kommt es zum Energietransfer in den Atlantik-Europa Bereich. Dabei wird eine Struktur aufgebaut, die Ähnlichkeit mit einem Pazifik-Nordamerika (PNA) Muster besitzt (Abbildung 5c). Nach 192 Stunden befindet sich schließlich ein

Teil der Störungsamplitude im Zielgebiet. Wie man in Abbildung 5d erkennen kann, ist auch noch in anderen Gebieten Störungsamplitude vorhanden. Die lokale Optimierung schließt Störungen außerhalb des Zielgebietes nicht aus. Der Anteil, welcher sich im Nordatlantik-Europa Bereich befindet, repräsentiert die maximal mögliche kinetische Energie. Mehr ist für dieses Gebiet, unter dem verwendeten Grundstrom und Optimierungszeit, nicht erreichbar.

Zum Schluß soll noch darauf hingewiesen werden, daß das gerade besprochene keine universelle Gültigkeit besitzt. D.h., es soll nicht bedeuten, daß ausschließlich PNA ähnliche Muster die kinetische Energie vom Pazifik in den Atlantik transportieren. Es erfolgte vielmehr die Beobachtung verschiedenster Prozesse, die alle davon abhängen welche Gestalt der zugrundeliegende Grundstrom besitzt. Je nach Grundstrom, kommt es zu einer Bevorzugung des einen oder anderen von den vier möglichen Startgebieten. Zum Beispiel, dominieren bei einem stark ausgeprägten Atlantikjet Störungen, die sich in dieser Region befinden. Sie wandern dann auf direktem Weg in das Zielgebiet hinein. Dabei durchlaufen sie die oben beschriebenen Prozesse für barotropes Störungswachstum. Anders verhält es sich bei der Anwesenheit von Blockierenden Hochs über dem Atlantik. Hier dominieren ortsfeste Strukturen in diesem Bereich, die nach Ablauf des Optimierungsintervalls große Ähnlichkeit mit der von Metz [9] berechneten Blocking-Anomalie besitzen.

5 Diskussion und Schlußfolgerung

Interessant ist die Ähnlichkeit von EOF's und optimalen Vektoren auf der Südhalbkugel. Dies ist um so erstaunlicher, da die optimalen Vektoren nur für ein barotropes Modell berechnet worden. Damit können sie natürlich nur den barotropen Anteil des Störungswachstums widerspiegeln. Sie geben Auskunft über theoretisch mögliches Störungswachstum. Die EOF's dagegen stellen Beobachtungswerte, in Form von Varianzmustern, über die wandernden Hoch- und Tiefdruckgebiete dar. Für sie sind barokline Prozesse dominant. Um die Frage zu verfolgen, woher die Gemeinsamkeiten in Gestalt und Lage zwischen den EOF Mustern und den optimalen Vektoren herrühren, sind weitere Untersuchungen nötig. Nützlich dafür ist wahrscheinlich die Berechnung von baroklinen optimalen Vektoren für die Südhalbkugel.

Weiterhin kam der Nordatlantik-Europa Raum als Optimierungsgebiet zum Einsatz. Hier sind die unterschiedlichsten Grundströme mittels lokaler optimaler Vektoren untersucht worden. Aus der Art der Optimierung ergibt sich, daß die Anfangsstörungen nicht im Zielgebiet liegen müssen. Auf diese Weise können die optimalen Vektoren als Energiequelle neben dem Atlantikjet auch den Asiatischen Subtropenjet für ihr Störungswachstum nutzen. Dieser Sachverhalt trifft aber nur für lange Optimierungszeiten zu. Für Zeiten kleiner 192 Stunden befinden sich die Anfangsstörungen hauptsächlich in der Atlantikjetregion, da es innerhalb dieser kurzen Intervalle nicht möglich ist, die Energie vom pazifischen- in den atlantischen Bereich zu transportieren.

Aus der Art und Weise, wie die Störungen in das Zielgebiet wandern, ist zu er-

kennen, welche Gebiete miteinander in Wechselwirkung stehen. Teleconnections zeigen solche Fernwirkungen an, Wallace u. Gutzler [14]. Ihre Berechnung erfolgt über statistische Methoden. Ein Vertreter dieser Strukturen ist das sogenannte PNA Muster. Es repräsentiert die Wechselwirkung zwischen Pazifik und Nordamerika auf der Nordhalbkugel. Dabei handelt es sich um eine ortsfeste niederfrequente Oszillation der Atmosphäre.

Bei ihrer Wanderung in den Nordatlantik-Europa Raum kann es vorkommen, daß die optimalen Vektoren PNA ähnliche Muster anregen oder solche Strukturen nutzen, um die kinetische Energie in das Zielgebiet zu transportieren. Dieses Verhalten wird dabei maßgeblich vom zugrundeliegenden Grundstrom bestimmt. Neben diesem PNA Transport konnten die unterschiedlichsten Mechanismen beobachtet werden. Dabei hat es den Anschein, daß die Strukturen, die sich im pazifischen Raum entwickeln, als Auslöser für die Atlantik-Strukturen fungieren.

Die Frage des Einflusses des jeweiligen Grundstromes auf den Energietransport ist noch genauer zu untersuchen. Da hinein spielt auch die Frage welche Bedingungen erfüllt sein müssen, damit solch ein PNA Muster angeregt wird. Auch die anderen Transportmechanismen sollten eine genauere Prüfung erfahren. Vielleicht lassen sich in ihnen Gemeinsamkeiten erkennen, die auf andere Wechselwirkungen zurückzuführen sind.

6 Anhang

6.1 Selbstadjungiertheit des Laplaceoperators

Aus der Struktur des Laplaceoperators in Kugelkoordinaten und der Gestalt der Kugelflächenfunktionen folgt $\nabla^2 Y_{m,n}(\mu, \lambda) = \nabla^2 Y_{m,n}^*(\mu, \lambda)$ und damit $\nabla^2 \psi(\mu, \lambda, t) = \nabla^2 \psi^*(\mu, \lambda, t)$, woraus die Selbstadjungiertheit von ∇^2 folgt, d.h.: $\{\psi, \nabla^2 \psi\} = \{\nabla^2 \psi, \psi\}$. Daraus folgt weiterhin, daß die kinetische Energie reell ist

$$E_{kin} = -\{\psi, \nabla^2 \psi\} = -\{\nabla^2 \psi, \psi\}^* = -\{\psi, \nabla^2 \psi\}^* = E_{kin}^*$$

6.2 Projektionsoperator

Bei Verwendung des Euklidischen Skalarproduktes ergibt sich für den Projektionsoperator folgende Matrixdarstellung Buizza [3]

$$\mathbf{T}_g = \mathbf{S}^{-1} \mathbf{G} \mathbf{S}$$

wobei die Matrizen auf der rechten Seite folgende Bedeutung besitzen:

1. \mathbf{S} führt die Transformation Spektral \rightarrow Gitter durch, d.h.: der k-dimensionale Vektor Ψ wird in den l-dimensionalen Vektor Ψ_g überführt. k ist dabei die Anzahl der Kugelfunktionskoeffizienten und l die Anzahl der Gitterpunktwerte. \mathbf{S} ist eine lineare Abbildung vom \mathcal{R}^k in den \mathcal{R}^l , sie wird durch eine $(l \times k)$ -Matrix repräsentiert.

2. \mathbf{G} legt das Zielgebiet g auf dem Gitter fest, d.h.:

$$\mathbf{G}\Psi_g = \begin{cases} \Psi_g(\mu_i, \lambda_j) & (\mu_i, \lambda_j) \in g \\ 0 & (\mu_i, \lambda_j) \notin g \end{cases}$$

Dieser Operator wird durch eine $(l \times l)$ -Diagonalmatrix repräsentiert.

3. \mathbf{S}^{-1} führt die Rücktransformation Gitter \rightarrow Spektral durch, welche durch eine $(k \times l)$ -Matrix dargestellt wird.

Literatur

- [1] J. Barkmeijer. Local error growth in a barotropic model. *Tellus 44A*, 314-323, 1992.
- [2] M. D. Borges and D. L. Hartmann. Barotropic instability and optimal perturbations of observed nonzonal flows. *J. Atmos. Sci.* 49, 335-354, 1992.
- [3] R. Buizza. Localization of optimal perturbations using a projection operator. *Q. J. R. Meteorol. Soc.* 120, 1647-1681, 1994.
- [4] R. Buizza, J. Tribbia, F. Molteni, and T. Palmer. Computation of optimal unstable structures for a numerical weather prediction model. *Tellus 45A*, 388-407, 1993.
- [5] B. Farrell. Optimal excitation of neutral rossby waves. *J. Atmos. Sci.* 45, 163-172, 1988.
- [6] H. Heinrich. Optimal wachsende Störungen in einem barotropen Modell. *Diplomarbeit am Institut fuer Meteorologie der Universitaet Leipzig*, 1999.
- [7] E. N. Lorenz. Deterministic non-periodic flow. *J. Atmos. Sci.* 20, 130-141, 1963.
- [8] E. N. Lorenz. A study of the predictability of a 28-variable atmospheric model. *Tellus 17*, 321-333, 1965.
- [9] W. Metz. Optimal relationship of large-scale flow patterns and the barotropic feedback due to high-frequency eddies. *J. Atmos. Sci.* 48, 1141-1159, 1991.
- [10] W. Metz. Singular modes and low-frequency atmospheric variability. *J. Atmos. Sci.* 51, 1740-1753, 1994.
- [11] F. Molteni and T. Palmer. Predictability and finite-time instability of the northern winter circulation. *Q. J. R. Meteorol. Soc.* 118, 269-298, 1993.
- [12] J. Oortwijn. Growth properties of optimal transition perturbations. *J. Atmos. Sci.* 56, 2491-2511, 1999.

- [13] F. Selten. An efficient description of the dynamics of barotropic flow. *J. Atmos. Sci.* 52, 915-936, 1995.
- [14] J. Wallace and D. Gutzler. Teleconnections in the geopotential height field during the northern hemisphere winter. *Mon. Wea. Rev.* 109, 784-812, 1981.

Adresse der Autoren:

Harald Heinrich, Werner Metz
Universität Leipzig, Institut für Meteorologie
Stephanstr. 3
04103 Leipzig

Singular vectors of Eady-models with $\beta \neq 0$ and $q' = 0$

Robin Faulwetter and Werner Metz

Summary:

As pointed out by Farrell, a normalmode analysis alone may be not enough for a convincing investigation of baroclinic stability. In some models growth rates can be achieved large enough to enable nonlinear growth also in parameter ranges of neutral normalmodes. According to Farrell one has also to consider that structures, which achieve optimal growth for a given, fixed time interval (i.e. the singular vectors).

Fischer (1998) investigated this problem for the classical Eady-model with $q' = 0$ - a case which can be treated analytically. In this paper we want to give a short overview of an investigation of singular vectors in Eady-models with $\beta \neq 0$ and $q' = 0$. Our aim was to understand the influence of $\beta \neq 0$ on optimal growth.

Qualitative differences to Fischer's results are only found at small wavenumbers below the longwave cutoff. The most remarkable difference beyond the longwave cutoff is the fact, that the singular vectors of the model with $\beta \neq 0$ grow faster in the upper half of the fluid than in the lower half for the considered basic flows. The growth rates for parameter ranges of neutral normalmodes are too small to enable nonlinear growth effects in meteorologically relevant times.

For long timescales we find, that the cutoffs must be understood more as a smooth transition to instability.

1 Introduction

Theories of baroclinic instability of synoptic and planetary scale flows are based upon the pioneering work of Charney (1947) and Eady (1949), who used the well known normalmode analysis to analyze their models. Eady (1949) found two discrete normalmodes which are not neutral - a decaying one and a growing one, whose phase speed, growth rate and wavelength were in good agreement with that of observed cyclones. But there were discrepancies between the structure of the unstable normalmodes, which are separable in time and space and the structures of observed cyclones, which are not separable.

Pedlosky (1964) treated the Eady - model as an initial value problem and pointed out, that the discrete normalmodes of the Eady - model are not complete in the sense that an arbitrary disturbance can not be uniquely expressed in terms of the discrete normalmodes. He pointed out that the discrete normalmodes must be supplemented by a continuous spectrum of stable modes to obtain a complete set of modes. Following Eady and Charney these stable modes have long been neglected because of their stability.

Farrell (1982) used this fact to overcome some discrepancies between observations and baroclinic theory, which was governed by the normalmode paradigm up to that point. He computed the growth of disturbances in different baroclinic models and found, that the Eady - initial value problem has solutions which are growing much faster than the unstable Eady - mode, when the continuous spectrum was included. In addition he showed that disturbances in the Eady - model can extract large amounts of energy from the background flow also for

parameter ranges, where no unstable modes exist. Lindzen et al. (1982) pointed out to this fact, too. Farrell (1984) considered the evolution of disturbances in the Eady - model, which consisted of the two discrete modes and a contribution of the continuous spectrum. He showed that rapid cyclogenesis and Petterssen - type - B - cyclogenesis can be understood in terms of this initial - value problem.

Rotunno and Fantini (1989) formed combinations of the discrete normalmodes and showed that these disturbances can also extract large amounts of energy from the background flow by constructive interference. They showed that rapid cyclogenesis and Petterssen - type - B cyclogenesis can also be understood without consideration of the continuous spectrum.

According to Farrell (1988) there are two interesting initial value problems in baroclinic theory:

First, one can search that disturbance, which deposits the maximal energy into a given stable or unstable mode in a given time (namely the adjoint mode; see chapter 4 in Farrell (1988)).

Second, one can search that disturbance, which achieves the optimal growth in a given finite time, where one has to specify, how the growth of a disturbance shall be measured. (We will do this in chapter 2.3.)

The second initial value problem is the crucial one for the short and middle term predictability of baroclinic flows. If an optimal disturbance is able to become large enough to enable nonlinear growth, one cannot say that the flow is stable although a normalmode analysis might yield stability.

Anyway it is clear that in a meteorological context disturbances, which attain the maximal growth in timescales, which are known to be significant for cyclogenesis, are much more interesting than unstable normalmodes.

In the following we will call disturbances, which achieve the maximal growth in a given finite time, “singular vectors”.¹

Fischer (1998, in the following F98) investigates the singular vectors of the classical Eady model with perturbation potential vorticity $q' = 0$ ² analytically. In this paper, we reexamine Fischer’s work and present some extensions to it. In particular, we will incorporate the modifications of the Eady model (i.e. $\beta \neq 0$), introduced by Lindzen (1994, in the following L94) into the singular vector problem of Fischer.

2 Theory

2.1 Formulation of the models

For the convenience of the reader we present the basic equations of Eady - type - models and briefly recapitulate Lindzen’s ideas to modify Eady’s model with $\beta \neq 0$.³

The basic equation is the nondimensional linearized quasigeostrophic equation for conservation of potential vorticity (for details see Lindzen (1990); Pedlosky (1987)):

¹We will recapitulate the precise definition of singular vectors in chapter 2.2

²As we will see in the next chapter the additional constraint $q' = 0$ is equivalent to the neglect of the continuous spectrum.

³We only treat the 2d - case but wish to point out the reader that there are no qualitative differences between the results for the 2d- and the 3d-case. The 2d - model can be imagined as a fluid which has an infinite extension in y-direction or alternatively no extension in y-direction at all.

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)q' + \frac{\partial\phi}{\partial x}\frac{\partial\bar{q}}{\partial y} = 0, \quad (1)$$

where ϕ is the perturbation streamfunction. q' and \bar{q} are given by:

$$q' = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial}{\partial z}\left(\frac{1}{S}\frac{\partial\phi}{\partial z}\right) \quad (2)$$

$$\frac{\partial\bar{q}}{\partial y} = \beta - \frac{\partial}{\partial z}\left(\frac{1}{S}\frac{\partial U}{\partial z}\right) = 0, \quad (3)$$

where we assumed $\rho_S(z) = \text{const.}$. The last identity is the well known Eady - condition. S is a function of z in general:

$$S = S(z) = \frac{N(z)^2 D^2}{f_0^2 L^2}, \quad (4)$$

where D is the distance between the bottom and the upper rigid lid. In the following $\frac{\partial U}{\partial z}$ will be abbreviated as m . Using the quasigeostrophic form of the first law of thermodynamics and the thermal wind relation we obtain the following linearized boundary conditions at $z = z_B, z_T$:

$$\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\frac{\partial\phi}{\partial z} - m\frac{\partial\phi}{\partial x} = 0. \quad (5)$$

If we make a normalmode ansatz for the perturbation streamfunction: $\phi(x, z, t) = A(z)e^{ik(x-ct)}$, (1,2) can be simplified with (3):

$$(U(z) - c)\left(\frac{\partial}{\partial z}\left(\frac{1}{S}\frac{\partial A}{\partial z}\right) - k^2 A\right) = 0. \quad (6)$$

Solutions of this equation must satisfy:

$$\frac{\partial}{\partial z}\left(\frac{1}{S}\frac{\partial A}{\partial z}\right) - k^2 A = \widehat{D}\delta(U(z) - c) \quad \widehat{D} \neq 0 \quad (7)$$

for fixed c .

If the right hand side of equation (7) is zero for all $z \in [z_B, z_T]$ i.e. $c < U_{min}$ or $c > U_{max}$, we obtain in general two solutions depending on $S = S(z)$, With the aid of the boundary conditions one can obtain the discrete normalmodes and the corresponding phase speeds.

On the other hand, if $c = U(z_C)$ for $z_C \in (z_B, z_T)$, solutions exist for every value of z_C . These solutions are called the continuous spectrum. The continuous spectrum renders the problem of finding the singular vectors very difficult.

F98 replaced equations (1) and (2) with the constraint $q' = 0$. It is evident from (1) to (3) and (6), that this assumption is equivalent to a restriction on the discrete normalmodes.

In this work we follow F98 by assuming $q' = 0$:

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial}{\partial z}\left(\frac{1}{S}\frac{\partial\phi}{\partial z}\right) = 0. \quad (8)$$

We apply periodic boundary conditions in x-direction, which allows us to write:

$$\phi(x, z, t) = \text{Re}\left(\sum \varphi_k(z, t)e^{ikx}\right) \quad (9)$$

(8) and (9) yield:

$$-\mu^2 \varphi_k(z, t) + \frac{\partial}{\partial z} \left(\frac{1}{\hat{S}} \frac{\partial \varphi_k(z, t)}{\partial z} \right) = 0 \quad (10)$$

$$\mu := \sqrt{S_0} k \quad \hat{S}(z) := \frac{S(z)}{S_0} \quad S_0 = S(z_B) \quad (11)$$

In general (10) and (8) have solutions of the form:

$$\varphi_k(z, t) = a_k(t) f_k(z) + b_k(t) g_k(z) \quad (12)$$

$$\Rightarrow \phi(x, z, t) = \text{Re} \left(\sum_{K=1}^{\infty} (a_k(t) f_k(z) + b_k(t) g_k(z)) e^{ikx} \right) \quad (13)$$

$$\alpha_k(x, z) := f_k(z) e^{ikx} \quad \beta_k(x, z) := g_k(z) e^{ikx}. \quad (14)$$

The time dependency of the a_k 's and b_k 's can be obtained from the boundary conditions (5):

$$\frac{1}{k} \frac{\partial}{\partial t} \frac{\partial}{\partial z} \varphi_k(z, t) + iU \frac{\partial}{\partial z} \varphi_k(z, t) - im(z) \varphi_k(z, t) = 0 \quad z = z_B, z_T. \quad (15)$$

For each wavenumber k this gives two equations for the two unknowns a_k and b_k . According to (10) and (11) the solution structure (in space) depends on the parameter μ and the structure of the the normalized stability parameter \hat{S} . We will see later that it is advantageous to treat the system in terms of the continuous parameter μ and not in terms of the discrete parameter k . For this reason it is necessary to rescale the time in order to eliminate k from (15) by introducing a new dimensionless time \hat{t} :

$$t^* = \frac{L}{U} t = \frac{L\sqrt{S_0}}{U} \hat{t} \quad (16)$$

Using (4) and (11) the boundary conditions (15) become:

$$\frac{1}{\mu} \frac{\partial}{\partial \hat{t}} \frac{\partial}{\partial z} \varphi_k(z, t) + iU \frac{\partial}{\partial z} \varphi_k(z, t) - im(z) \varphi_k(z, t) = 0 \quad z = z_B, z_T. \quad (17)$$

In the following we omit the index “ k ” and write t instead of \hat{t} . With $U(z = z_B) = 0$ and $U(z = z_T) = 1$, we obtain the governing equations of the system's dynamics:

$$\left(\frac{\partial}{\partial t} + i\mu \right) \frac{\partial \varphi(z, t)}{\partial z} - i\mu m \varphi(z, t) = 0 \quad z = z_T \quad (18)$$

$$\frac{\partial^2 \varphi(z, t)}{\partial z \partial t} - i\mu m \varphi(z, t) = 0 \quad z = z_B. \quad (19)$$

L94 supposed $\beta \neq 0$ and satisfied (3) by an adjustment of $U(z)$ and $S(z)$. Equation (3) yields:

$$\frac{m(z)}{S(z)} = \beta z + \frac{m_0}{S_0}, \quad (20)$$

where $m_0 = m(z_B)$ must be determined by the boundary conditions for U , and $S_0 = S(z_B)$ is given by (4). L94 examined two cases:

1) $S = S_0 = \text{const.}$. The determination of $U(z)$ from equation (20) requires the specification of the vertical coordinate:

$$z_B = -\frac{1}{2} \quad z_T = \frac{1}{2}$$

The integration of (20) results in:

$$U(z) = \frac{B}{2} \left(z^2 - \frac{1}{4} \right) + z + \frac{1}{2} \quad B := \beta S_0. \quad (21)$$

In this case the solutions must satisfy:

$$-\mu^2 \varphi(z, t) + \frac{\partial^2 \varphi(z, t)}{\partial z^2} = 0 \quad (22)$$

$$\Rightarrow \varphi(z, t) = a(t) \sinh(\mu z) + b(t) \cosh(\mu z) \quad (23)$$

$$\Rightarrow \alpha_k(x, z) = \sinh(\mu z) e^{ikx} \quad \beta_k = \cosh(\mu z) e^{ikx}. \quad (24)$$

This is the solution for the classical Eady-model, too. Please note, that (21) delivers the result for the classical Eady model, when B is set to zero. For this reason we can treat the classical model as a special case of Lindzen's " $S = \text{const.}$ -model".

2) The second case investigated by Lindzen is given by:

$$S(z) = \frac{S_0}{1 + \delta \frac{B}{m_0} z} \quad \delta \in (0, 1] \quad (25)$$

$$\Rightarrow m(z) = m_0 \frac{1 + \frac{B}{m_0} z}{1 + \delta \frac{B}{m_0} z}. \quad (26)$$

We investigate this case in a slightly different coordinate system:

$$z_B = 0 \quad z_T = 1.$$

Though it is easily possible to integrate (26), it is not possible to determine the constants produced by integrations analytically with the usual boundary conditions $U(0) = 0$ and $U(1) = 1$. As the formulas obtained for the calculation of the singular vectors are so long and complicated, that they don't allow any insight, we solved them with Maple. For that reason we decided to use Maple tools too for the numerical adjustment of m_0 to $U(1) = 1$, in order to preserve the comparability of the results. Please note that once the appropriate value of m_0 was given, all calculations were carried through analytically. Please note that m_0 is monotonically growing with δ and that $m(1)$ is monotonically decreasing with δ , what is of importance for the dynamics according to (20) and the fact, that the local efficiency of energy conversion from the basic flow to disturbance energy is given by $\frac{m}{S}$. In this case Lindzen solved the problem (10) with modified Bessel functions of zeroth order:

$$\varphi_k(z, t) = a(t) I_0(u(z)) + b(t) K_0(u(z)) \quad (27)$$

$$u(z) = \frac{2\mu m_0}{\delta B \sqrt{\hat{S}(z)}} = \frac{2\mu m_0}{\delta B} \sqrt{1 + \delta \frac{B}{m_0} z} \quad \delta \in (0, 1].$$

If negative values of z were allowed, one might obtain imaginary values of $u(z)$, what would lead to computational problems. So we chose the coordinate system as explained before. In

agreement with the results for the $S = \text{const.}$ - model a normalmode analysis of this model, which will be called the $S = S(z)$ - model in the following, yields a longwave-cutoff for baroclinic instability, too (see Lindzens fig 3). For short waves there are only very small differences to the results obtained for the classical model.

2.2 Singular vectors

Up to this point we did not say, how the growth of disturbances is measured. We define the growth rate α :

$$\alpha = \frac{\|\phi(t)\|}{\|\phi(t_0)\|} = \frac{\|\phi(t_0 + T)\|}{\|\phi(t_0)\|}, \quad (28)$$

where $\|\cdot\|$ should be a physically relevant norm, which must be given in terms of an inner product $\langle \cdot, \cdot \rangle$. ϕ is the disturbance streamfunction. The aim is to find that disturbance ϕ , which optimizes the growth rate α with fixed T . In the following T will be called the optimization time and t_0 is set to zero without loss of generality.

Joly and Thorpe (1991) and Joly (1995) formulate this optimal growth problem in a more compact form:

The propagator $M(T)$ of a linear system relates the disturbance at time T to the initial disturbance:

$$\phi(T) = M(T)\phi(0). \quad (29)$$

For convenience the propagator $M(T)$ is abbreviated as M . The optimal perturbation is found as that eigenvector of M^*M which has the largest eigenvalue. M^* is the adjoint of M with respect to the chosen inner product. Usually these eigenvectors are called ‘‘singular vectors’’ and are ordered with decreasing magnitude of the eigenvalues.

Now suppose that all solutions of the model can be expressed as a linear combination of modes, and let \underline{B} be the basis of the vector space of all solutions, i.e. in our case $\phi(x, z, t) = \psi(t) \cdot \underline{B}(x, z)$, where $\psi(t)$ is a row vector and \underline{B} is a column vector. Then one can formulate a propagator equation in the vector space \underline{B} :

$$\psi(T)^T = M\psi_0^T. \quad (30)$$

The inner product $\langle \cdot, \cdot \rangle$ in the vector space \underline{B} is given by a metric matrix Σ :

$$\langle \phi, \phi \rangle = \bar{\psi}\Sigma\psi^T \quad (31)$$

With the aid of (30) and (31) equation (28) yields:

$$\alpha = \frac{\|\phi(T)\|}{\|\phi(t_0)\|} = \frac{\langle \phi(T), \phi(T) \rangle}{\langle \phi(t_0), \phi(t_0) \rangle} = \frac{\overline{\psi_0 M^T \Sigma M \psi_0^T}}{\overline{\psi_0 \Sigma \psi_0^T}} = \frac{\overline{\psi_0 \Sigma \Sigma^{-1} \overline{M^T} \Sigma M \psi_0^T}}{\overline{\psi_0 \Sigma \psi_0^T}} \quad (32)$$

(See also Heinrich (1999).) We see that according to the definition of the adjoint matrix with respect to an inner product M^* must be identical to $\Sigma^{-1} \overline{M^T} \Sigma$. Therefore the singular vectors can be calculated as the eigenvectors of $M^*M = \Sigma^{-1} \overline{M^T} \Sigma M$; the appropriate growth rates are given by the accompanying eigenvalues.

In our case \underline{B} consists of the α_k 's and β_k 's introduced by (14), and $\psi(t)$ consists of the a_k 's and b_k 's introduced in (12). As pointed out before (15) yields two coupled equations for each wavenumber, against what the different wavenumbers are independent of each other.

Therefore the propagators (expressed in the vector space \underline{B}) of the models described in the preceding section have the following form:

$$M = \begin{pmatrix} M_1 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ & 0 & 0 & 0 & 0 & & & \\ & 0 & 0 & M_2 & 0 & 0 & \cdot & \cdot & \cdot \\ & 0 & 0 & 0 & 0 & & & & \\ & 0 & 0 & 0 & 0 & M_3 & & & \\ \cdot & & & & & & \cdot & & \\ \cdot & & & & & & & \cdot & \\ \cdot & & & & & & & & \cdot \end{pmatrix}. \quad (33)$$

The (2×2) -matrices M_K on the diagonal of (33) are obtained by inserting (12) into the two boundary conditions (15) and writing these two equations in matrix form.

Now one has to calculate the matrix Σ , which describes the chosen norm in terms of the coordinate vector ψ of a solution in the vector space \underline{B} . All norms, which are physically relevant, are given by an inner product, which involves an integration over the considered volume. For all these norms we obtain with the aid of (14):

$$\langle \alpha_k, \alpha_l \rangle = \langle \alpha_k, \beta_l \rangle = \langle \beta_k, \beta_l \rangle = 0 \quad \forall k \neq l. \quad (34)$$

Because of the bilinearity of inner products (31) and (13) yield the special form of the metric matrix Σ :

$$\Sigma = \begin{pmatrix} \langle \alpha_1, \alpha_1 \rangle & \langle \alpha_1, \beta_1 \rangle & & 0 & 0 & & \cdot & \cdot & \cdot \\ \langle \beta_1, \alpha_1 \rangle & \langle \beta_1, \beta_1 \rangle & & 0 & 0 & & & & \\ & 0 & 0 & \langle \alpha_2, \alpha_2 \rangle & \langle \alpha_2, \beta_2 \rangle & & & & \\ & 0 & 0 & \langle \beta_2, \alpha_2 \rangle & \langle \beta_2, \beta_2 \rangle & & & & \\ \cdot & & & & & & \cdot & & \\ \cdot & & & & & & & \cdot & \\ \cdot & & & & & & & & \cdot \end{pmatrix} \quad (35)$$

So we see that Σ has the same form as M . One can easily prove that also the matrix $M^*M = \Sigma^{-1}\overline{M}^T\Sigma M$ is of this form and that the eigenvectors of such a matrix can be obtained as the eigenvectors of the (2×2) -matrices on the diagonal. This fact enables us to consider only one wavenumber. The first singular vector can then be obtained by calculating the singular vectors for each wavenumber and then searching for that wavenumber, whose first singular vector has the largest eigenvalue. As we saw earlier, all models which are considered in this paper are not dependent on the absolute value of the discrete wavenumber k , but on the continuous wavenumber μ . For that reason we can suppress all indices, set $k = 2\pi$ and calculate only with the wavenumber μ . We will abbreviate the first singular vector (for one wavenumber) simply as the ‘‘singular vector’’, because in our context only the first singular vector is interesting, as the second singular vector is the optimally decaying one.

2.3 Inner products

We chose to consider the kinetic energy per unit volume (K), the enstrophy per unit volume (V), the L_2 -norm and the available potential energy per volume (A) as norms for the measurement of growth. They are defined as follows:

$$K(\phi) = \frac{\rho_0}{2LD} \int_{z_B^*}^{z_T^*} \int_0^L v^{*2} dx^* dz^* = \frac{\rho_0 U^2}{2} \int_{z_B}^{z_T} \int_0^1 \frac{\partial \phi}{\partial x} \overline{\frac{\partial \phi}{\partial x}} dx dz \quad (36)$$

$$V(\phi) = \frac{1}{2LD} \int_{z_B^*}^{z_T^*} \int_0^L \left(\frac{\partial v^*}{\partial x^*} \right)^2 dx^* dz^* = \frac{U^2}{2L^2} \int_{z_B}^{z_T} \int_0^1 \frac{\partial^2 \phi}{\partial x^2} \overline{\frac{\partial^2 \phi}{\partial x^2}} dx dz \quad (37)$$

$$\langle \phi, \phi \rangle_0 = \int_{z_B}^{z_T} \int_0^1 \phi \overline{\phi} dx dz = (L_2(\phi))^2 \quad (38)$$

$$A(\phi) = \frac{\rho_0}{2LD} \int_{z_B^*}^{z_T^*} \int_0^L \frac{g^2}{\theta^{*2}} \frac{\theta^{*2}}{N^2} dx^* dz^* = \frac{\rho_0 g^2}{2N^2} \int_{z_B}^{z_T} \int_0^1 \frac{\partial \phi}{\partial z} \overline{\frac{\partial \phi}{\partial z}} dx dz \quad (39)$$

According to the definition (28) of the growth rate α the factors in front of the integrals are irrelevant for the growth measurement and they must not be considered for the calculation of singular vectors. As a consequence of (35) and (14) only the inner products of functions of the form $\phi = \varphi(z, t)e^{2\pi i x}$ must be calculated. One can easily prove that the norms K , V and $\langle \cdot, \cdot \rangle_0$ deliver the same result apart from irrelevant factors which are constant for any fixed wavenumber. Therefore the growth measured by these norms is identical, and we only have to consider the growth measured by K and A .

Before we go over to the presentation and discussion of our results we wish to recapitulate the equation for the growth of disturbance energy:

$$\underbrace{\int_{z_B}^{z_T} \frac{m}{S} \overline{\frac{\partial \phi}{\partial z} \frac{\partial \phi}{\partial x}} dz}_{\frac{\partial \overline{E}}{\partial t}} = \underbrace{\int_{z_B}^{z_T} \frac{1}{2S} \frac{\partial \phi}{\partial t} \overline{\left(\frac{\partial \phi}{\partial z} \right)^2} dz}_{\frac{\partial \overline{A}}{\partial t}} + \underbrace{\int_{z_B}^{z_T} \overline{w_1 \frac{\partial \phi}{\partial z}} dz}_{\frac{\partial \overline{K}}{\partial t}}, \quad (40)$$

where overbars denote the zonal average. Please remember the fact, that a direct conversion of mean flow available potential energy to disturbance kinetic energy is not possible. Therefore the lhs term in (40) is always a gain or loss of \overline{A} and not of \overline{K} .

2.4 Results for the inner product matricizes and propagators

The calculation of the (2×2) -matricizes Σ_A and Σ_K , which describe the considered scalarproducts for one wavenumber is obtained by inserting (24) in (36) or (39) and delivers the following results for the $S = const.$ -model:

$$\Sigma_K = \begin{pmatrix} \sinh \mu - \mu & 0 \\ 0 & \sinh \mu + \mu \end{pmatrix} \quad \Sigma_A = \begin{pmatrix} \sinh \mu + \mu & 0 \\ 0 & \sinh \mu - \mu \end{pmatrix} \quad (41)$$

apart from irrelevant factors. The simplicity of this result was the reason for the choice of the coordinate system, in which this model was treated.

The corresponding results for the $S = S(z)$ -model are obtained straightforwardly by inserting $\alpha_k = I_0(u(z))e^{ikx}$ and $\beta_k = K_0(u(z))e^{ikx}$ in (36) or (39).

The propagator is calculated by inserting (23) respectively (27) in (18) and (19) and writing the two results in matrix form, what leads for the $S = const.$ -model to:

$$M = e^{-i\mu c_0 T} \begin{pmatrix} \cos \frac{\mu w T}{2} - i \frac{v_0}{w} \sin \frac{\mu w T}{2} & i \left(\frac{v_0^2}{w d} - \frac{w}{d} \right) \sin \frac{\mu w T}{2} \\ -i \frac{d}{w} \sin \frac{\mu w T}{2} & \cos \frac{\mu w T}{2} + i \frac{v_0}{w} \sin \frac{\mu w T}{2} \end{pmatrix} \quad (42)$$

$$w = \sqrt{1 - \frac{4}{\mu} \coth \mu + \frac{4}{\mu^2} + \frac{B(\coth^2 \mu - 1)}{\mu^2}} \quad v_0 = \frac{B \sqrt{\coth^2 \mu - 1}}{\mu} \quad (43)$$

$$d = \coth \frac{\mu}{2} - \frac{2}{\mu} \quad c_0 = \frac{1}{2} - \frac{B \coth \mu}{2\mu}. \quad (44)$$

The analytical calculation of the propagator for the $S = S(z)$ -model results in very long formulas, which do not allow any insight. So it does not make any sense to give it explicitly.

3 Results and discussion

In the following we often will abbreviate “normalmode” with “NM”, “singular vector with respect to A ” with “S1A” and “singular vector with respect to K ” with “S1K”. The optimization interval will often simply be called T . We computed the S1A- and S1K-structures and growth rates for $T = 6h$, $T = 48h$ and $T = 72h$ and the NM-structures and growth rates in order to compare them with each other.

In this paper we can only deliver a small overview over the results and discussion presented in Faulwetter (2000), but we will publish these results soon in more detail.

In all cases we adopted the following standard parameters from F98.

f	$10^{-4} s^{-1}$
β_0	$1.62 \cdot 10^{-11} s^{-1} m^{-1}$
D	$9000 m$
U	$29.43 m s^{-1}$
$N^2(z_B)$	$1.2753 \cdot 10^{-4} s^{-2}$

Table 1: *Standard values of parameters.*

3.1 Singular vectors with respect to A

3.1.1 Growth rates

The S1A-growth rates are presented in figure 1. For all cases the curves are completely smooth in contrast to the NM-curves, which are not smooth at the points of the longwave and short-wave cutoff. The curves for S1A-growth rates have their maxima at wavenumbers $\mu \approx 1$ for all models with $\beta \neq 0$, corresponding to wavelengths of about 6000 km. The normalmodes achieve the greatest growth rates at wavelengths of about 4000 km.

Table 2 gives the values for the maxima of the curves for S1A- and NM-growth rates.

	$S = const.$			$S = S(z), \delta = 1.0$		
$T [h]$	S1A	NM	S1A/NM	S1A	NM	S1A/NM
48	214	17	12.59	357	30	11.90
72	515	75	6.87	1057	161	6.57
120	6170	1330	4.64	20000	4500	4.44
360	$3.05 \cdot 10^{11}$	$9.40 \cdot 10^{10}$	3.24	$8.54 \cdot 10^9$	$2.53 \cdot 10^9$	3.34
720	$2.67 \cdot 10^{22}$	$8.74 \cdot 10^{21}$	3.05	$2.00 \cdot 10^{19}$	$6.50 \cdot 10^{18}$	3.08

Table 2: *Comparison of S1A- and NM-growth rates.*

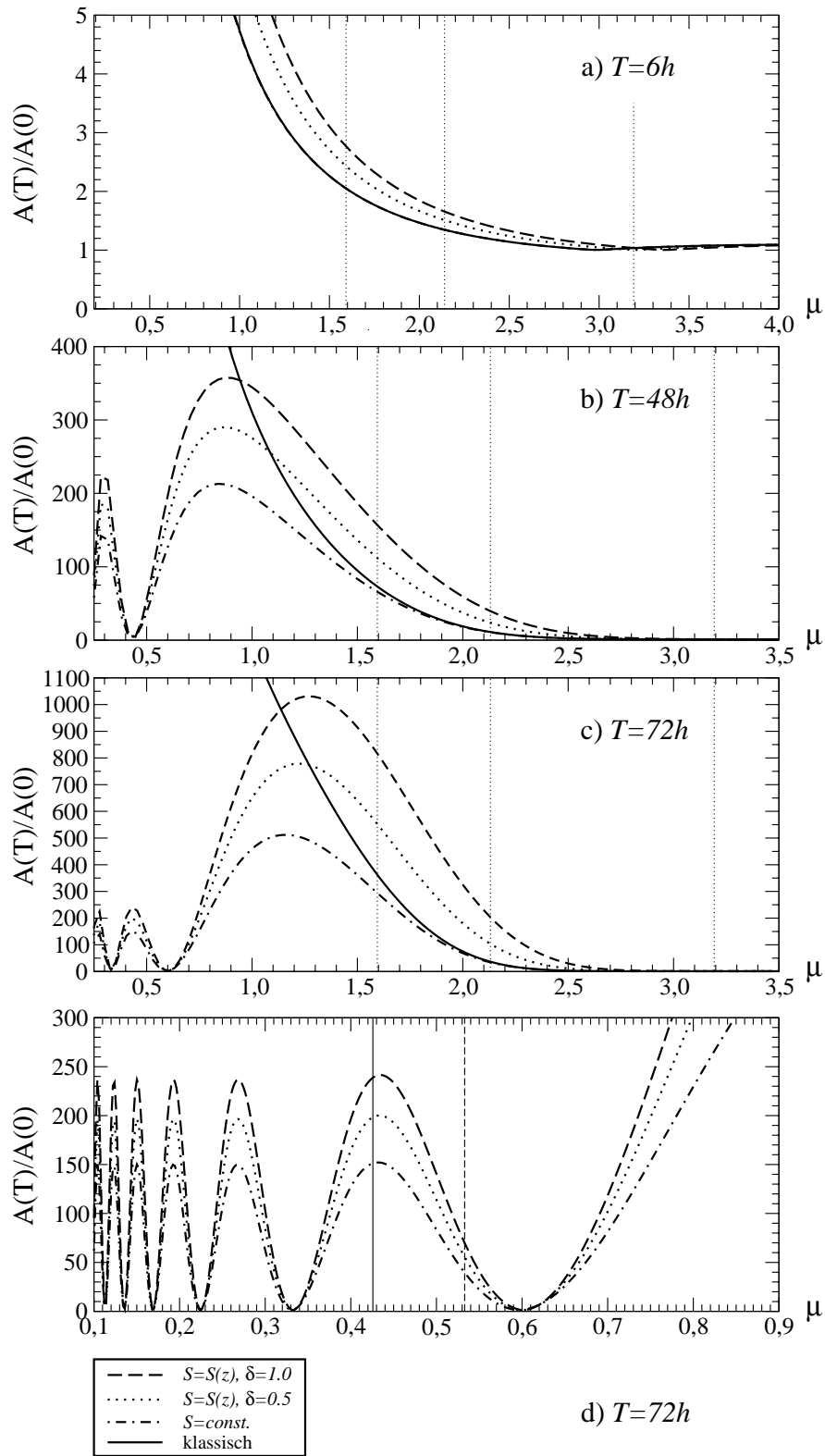


Figure 1: SIA - growth rates for $T = 6h$, $T = 48h$ and $T = 72h$.

For all optimization intervalls between $48h$ and $720h$ the singular vectors achieve much greater growth rates than the normalmodes. This fact is astonishing, because according to the normalmode - paradigm the unstable normalmode is expected to become the only important mode for very long times.

The behaviour of the S1A-growth rates is most interesting for small wavenumbers. F98 found, that the S1A-growth rates of the classical model go to infinity like $\frac{12f^2U^2T^{*2}}{D^2N^2\mu^2}$ when $\mu \rightarrow 0$ for all T . In addition to Fischers results we tried to gain some more physical insight in the behaviour of these very fast growing structures, with the aid of a Taylor - expansion for small wavenumbers. We found that the explosive growth is based on the fact $m(1) = m(0)$. Any difference between $m(0)$ and $m(1)$ will destroy the possibility of unbounded growth for small wavenumbers.

For that reason the S1A-growth rates of the models with $\beta \neq 0$ show a completely different behaviour below the longwave - cutoff: they oscillate with large amplitudes.

Due to the longwave - cutoff the S1A-vectors are only a linear combination of neutral modes, and therefore they can't achieve unbounded growth - rates for $\mu \rightarrow 0$. Of course any linear combination of neutral modes must be periodic in time. In the case of two modes the period \tilde{t} is given by:

$$\tilde{t} = \frac{2\pi}{\mu (Re(c_1) - Re(c_2))}, \quad (45)$$

where c_1 and c_2 are the phase speeds corresponding to the not - rescaled time. As the difference in the phase speeds of the two involved neutral modes grows as $1/\mu^2$ for $\mu \rightarrow 0$, the period decreases with decreasing μ .

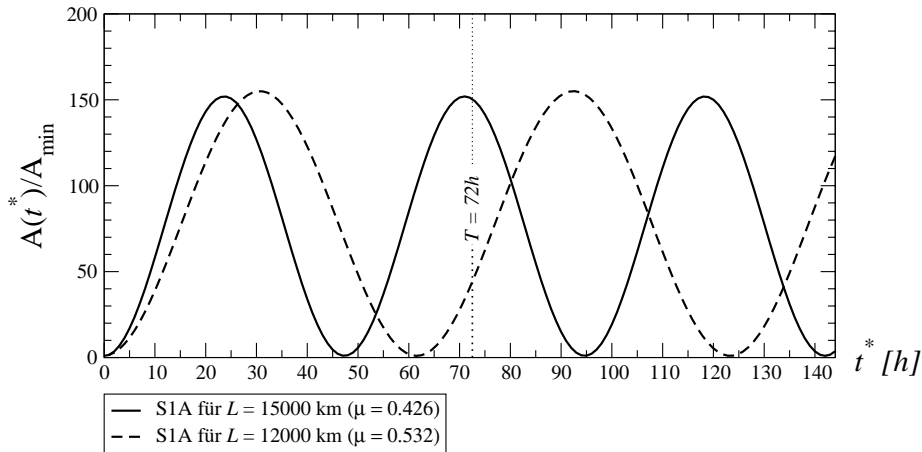


Figure 2: Time dependencies of the S1A - growth rates of the model with $S = const.$ for two different wavenumbers.

Consider now figure 2, which shows the time dependencies of the growth rates of the S1A-vectors for the $S = const.$ -model for two wavenumbers (which correspond to the vertical lines in figure 1d). For a fixed μ obviously all other linear combinations of the two neutral normalmodes than the singular vector have smaller amplitudes than the singular vector. For this reason, the singular vector is the same linear combination for all optimization intervalls and a fixed wavenumber. The problem of finding the optimally growing structure (i.e. the structure of the singular vector for $t = 0$) is therefore equivalent to finding the optimal initial point in the cycle shown in figure 2. Consider now a fixed T (e.g. $T = 72h$ corresponding to the vertical line in figure 2) and varying wavenumber μ . For $\mu = 0.532$ (dashed line) the

period of the growth of the optimal structure is only moderate in comparison to the growth at wavenumber $\mu = 0.426$ (solid line). The values of the growth rates at $t = T$ in figure 2 correspond to the points where the solid and dashed vertical lines in figure 1 intersects the growth rate curve for the considered model (dot-dashed line). When μ decreases furthermore there will soon be a wavenumber at which T is a multiple of the period of oscillation and no growth at all is possible. From this point of view the oscillations of the S1A-growth rates with μ become clear.

Remember now the fact, that all singular vectors have the maximal amplitude for a given, fixed μ . As pointed out before the problem of finding the optimally growing structure (i.e. the structure of the singular vector for $t = 0$) is therefore equivalent to the problem of finding the optimal initial point in a cycle like that shown in figure 2. For this reason the singular vectors must be the same linear combination of neutral modes for all T , and for all wavenumbers below the longwave cutoff there must be a maximal value of the S1A-growth rate which is independent of time. In the following the curve of these maximal values of S1 - growth rates for parameter ranges, where the normalmodes are neutral, will be called the “envelope”. It is possible to calculate envelopes analytically, but this is beyond the scope of this paper (for details and further discussion of the envelopes see Faulwetter (2000)).

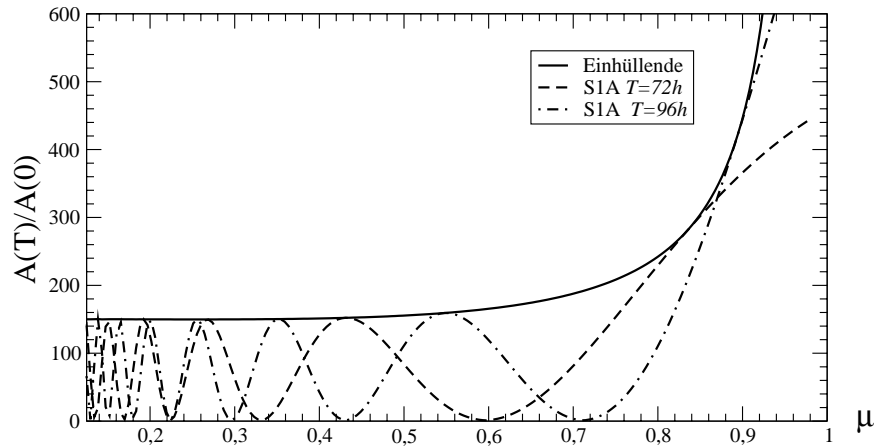


Figure 3: S1A - growth rates of the model with $S = const.$ for the optimization intervals $T = 72h$ and $T = 96h$ with the corresponding envelope for long waves.

Figure 3 shows the S1A-envelope for the $S = const.$ - model and low wavenumbers together with the S1A-growth rates for two different values of T . The envelope tends toward the value

$$1 + \frac{4}{B^2} \left(6 + \sqrt{36 + 3B^2} \right) \approx 1 + \frac{48}{B^2}$$

when $\mu \rightarrow 0$. The envelope has a minimum; in the special case it lies at $\mu = 0.255$ with a value of 149.8038. It remains nearly constant over a long range of wavenumbers. This behaviour can be found for all values of $B > 0$ and also for the corresponding envelope of the $S = S(z)$ - model.

The oscillations of the S1A-growth rates for short waves beyond the short wave cutoff have much smaller amplitudes. This can be understood when one remembers the fact, that the gain of disturbance available potential energy is proportional to the correlation of $\frac{\partial \phi}{\partial x}$ and $\frac{\partial \phi}{\partial z}$ (See (40)). The normalmodes have local extrema at the boundaries. It is clear from (24) that the extension of the disturbances in the vertical decreases with μ , and therefore also the correlation of $\frac{\partial \phi}{\partial x}$ and $\frac{\partial \phi}{\partial z}$ must be smaller for short waves than for long waves.

The fact that the S1A-growth rates for the $S = S(z)$ - model are larger than for the $S = \text{const.}$ - model becomes clear from equation (40). The local gain of A is proportional to $\frac{m}{S} = \beta z + \frac{m_0}{S_0}$, which is increasing monotonously with δ - due to the dependency of m_0 on δ .

3.1.2 Structures

Figure 4 shows the structures of the singular modes of the $S = S(z)$ - model for $t = 0$ and $t = T = 48h$ for the three wavelength $2000km$ ($\mu = 3.19$), $3000km$ ($\mu = 2.13$) and $4000km$ ($\mu = 1.6$)⁴, which are representative for short, middle and long wavelength. The results for the classical model and the $S = S(z)$ - model cannot be shown here; they can be found in Faulwetter (2000). The value of “d” in the plots gives the interval in which the isolines of disturbance streamfunction were plotted.

$L = 2000km$ (Figure 4a,b): For all models the saddle points, to which the disturbance is symmetric, is located exactly in the middle of the flow. The disturbance is confined to the boundaries and therefore the correlation between $\frac{\partial\phi}{\partial z}$ and $\frac{\partial\phi}{\partial x}$ remains weak and the growth rates small. The upper wave moves faster eastwards than the lower wave; when the minima (maxima) at the top make up the minima (maxima) at the bottom, A is growing until the minima (maxima) at the top are located exactly over the minima (maxima) at the bottom. Then disturbance energy is converted back into basic flow energy again, and so on. The difference between the models can only be seen in the difference between the phase speeds of the upper and lower wave.

$L = 3000km$ (Figure 4c,d): For the $S = \text{const.}$ - model and the classical model (not shown) the structures are similar to the NM-structures for this wavelength, in agreement with a point of contact between the S1A-growth rates and NM-growth rates found for all models at wavenumbers $\mu \approx 2$ ($L = 3000km$).

There is one qualitative difference between the models with $\beta \neq 0$ and the classical model. In the classical case, the saddle point remains exactly in the middle of the flow, whereas in the other cases this point moves from the upper into the lower part of the flow. It is remarkable, that the starting point of the saddle point descends with increasing δ . This behaviour was found in the whole parameter range of normalmode instability.

We want to deliver a plausible explanation for this behaviour: from equation (40) we see that the local “energy conversion efficiency” is equal to $\frac{m}{S} = \beta z + \frac{m_0}{S_0}$ and increases linearly with height. Therefore the disturbance grows faster at the top than at the bottom and the saddle point decreases with time. As m_0 increases with δ , the height - dependent part of the local “energy conversion efficiency” becomes less important with increasing δ . Therefore the relative difference between the local energy growth at the top and at the bottom decreases with increasing δ , and the saddle point descend slower.

$L = 4000km$ (Figure 4e,f): For all models the structure is nearly vertical for $t = 0$ and then obviously tends toward the NM-structure. The nearly vertical S1A-structure for $t = 0$ can be understood well by analysis of the Taylor - expansion of the model for low wavenumbers mentioned before (see Faulwetter (2000)).

$A(t = 0)$ is obviously small. Due to the advection of mean flow temperature over the whole height of the fluid the available potential energy can grow fast. It is clear, that in this way the growth rate $\frac{A(T)}{A(0)}$ is optimized.

In the case $T = 72h$ (not shown) the behaviour is very similar to that observed for the optimization interval $T = 48h$.

⁴See also the vertical lines in figure 1a-c.
The dimensional wavelength is given by $\Lambda = \frac{6385km}{\mu}$

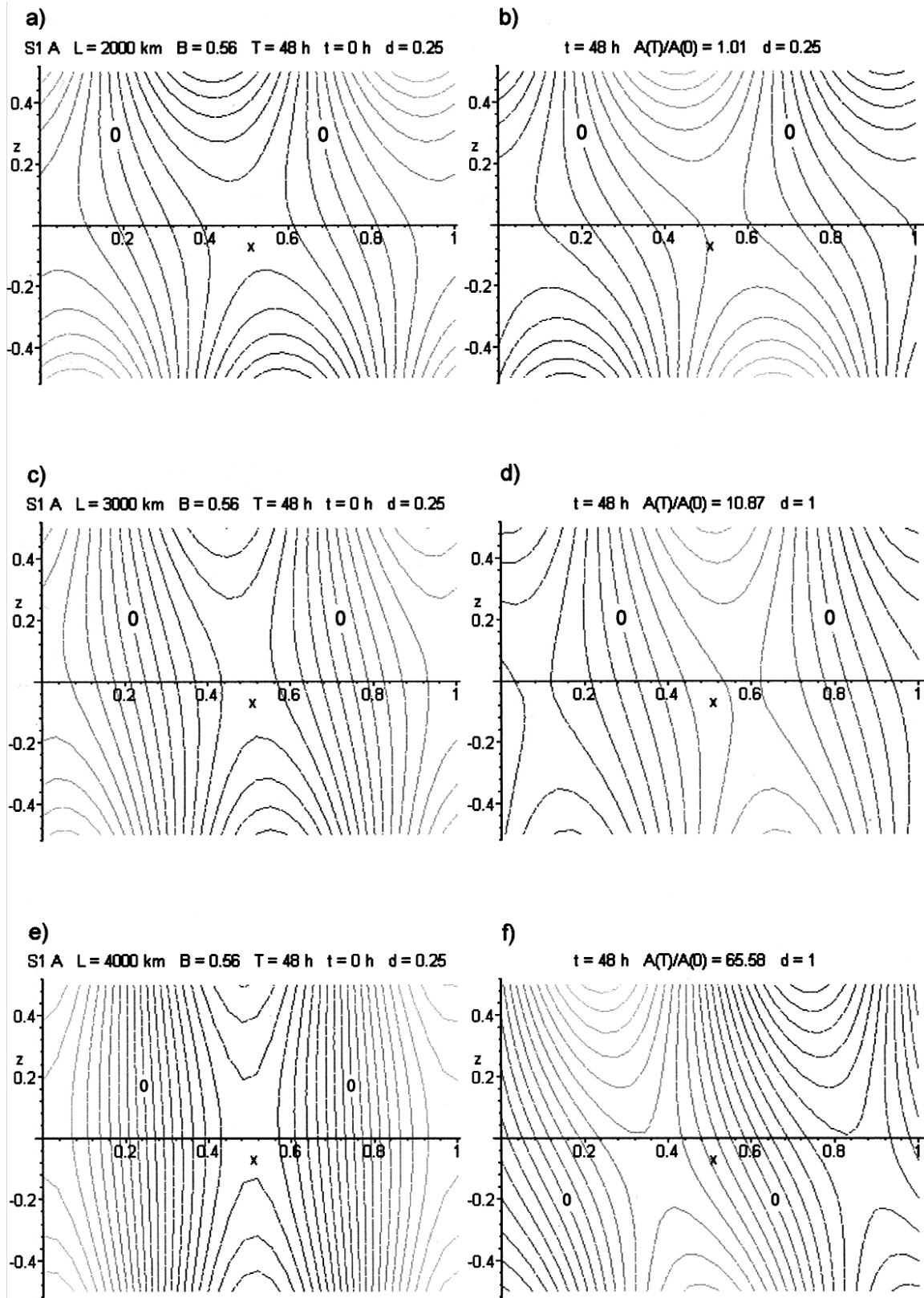


Figure 4: *SIA* - structures for $T = 48h$ and the wavelength $L = 2000km$, $L = 3000km$ and $L = 4000km$.

3.2 Singular vectors with respect to K

3.2.1 Growth rates

The S1K-growth rates are shown in figure 5. The S1K-growth rates have their maxima at wavenumbers $\mu \approx 2$ ($L \approx 3000km$). So the S1K-growth rates have their maxima at larger wavenumbers than the normalmodes - in contrast to the S1A-growth rates. As A-K-conversion is the only process which renders K we must conclude, that short waves are not as good as long waves in storing A and are much more efficient in A-K-conversion. We suggest the following physical explanation for this fact: following equation (40) A-K-conversion is equal to the integral of vertical temperature flux. As the vertical speed, which is proportional to vertical temperature flux, is of order $U \frac{D}{L}$, short waves are more effective in A-K-conversion than long waves - supposed that the correlation $\frac{w \frac{\partial \phi}{\partial z}}{\sigma(\frac{\partial \phi}{\partial z})\sigma(w)}$ is not changed too much.

We found that the S1K-growth rate maxima are approximately 30% larger than the corresponding NM-growth rates for values of T from $48h$ to $720h$.

In contrast to S1A-growth rates there are no points of contact between the S1K-growth rates and NM-growth rates - except of the local minima in the parameter ranges of neutral normalmodes.

The most remarkable difference to the S1A-growth rates is the fact that the S1K-growth rates go to one when $\mu \rightarrow 0$. It is possible to understand this fact with aid of the Taylor expansion of the basic functions in the classical case, too (see Faulwetter (2000)). But that discussion is too long for this paper.

The expansion of the full model equations yield the following results:

$$\frac{K(T)}{K(0)} = 1 + \frac{\mu T}{\sqrt{3}} + O(\mu^2)$$

in the classical case, and:

$$\frac{K(T)}{K(0)} = 1 + \frac{2}{\sqrt{3}B}\mu^2 + O(\mu^4)$$

for the envelope of the $S = const.$ -model. Below the envelope the S1K-growth rates oscillate in exactly the same way as the corresponding S1A-growth rates, i.e. with the same positions of extrema. This fact is evident from the discussion delivered in section 3.1.1.

The stabilizing influence of β decreases with increasing μ ; on the other hand the destabilizing effect of a stability function $S(z)$, which decreases with height, increases with increasing μ (not shown). For these reasons there are points of inflection between the S1K-growth rate curve of the classical model and the S1K-growth rate curves of the $S = S(z)$ -models at wavenumbers slightly beyond one; for short waves the $S = S(z)$ -model achieves greater S1K-growth rates than the other models

3.2.2 Structures

Figure 6 shows the structures of the singular modes of the $S = S(z)$ - model for $t = 0$ and $t = T = 48h$ for the three wavelength $2000km$ ($\mu = 3.19$), $3000km$ ($\mu = 2.13$) and $4000km$ ($\mu = 1.6$).

$L = 2000km$ (Figure 6a,b): As this wavenumber is in the range of neutral normalmodes, the principle of optimization is exactly the same as that explained in the previous section, and of course also the behaviour of these singular vectors is in principle the same as that of the corresponding S1A-vectors. The only model in which moderate growth of K can be achieved

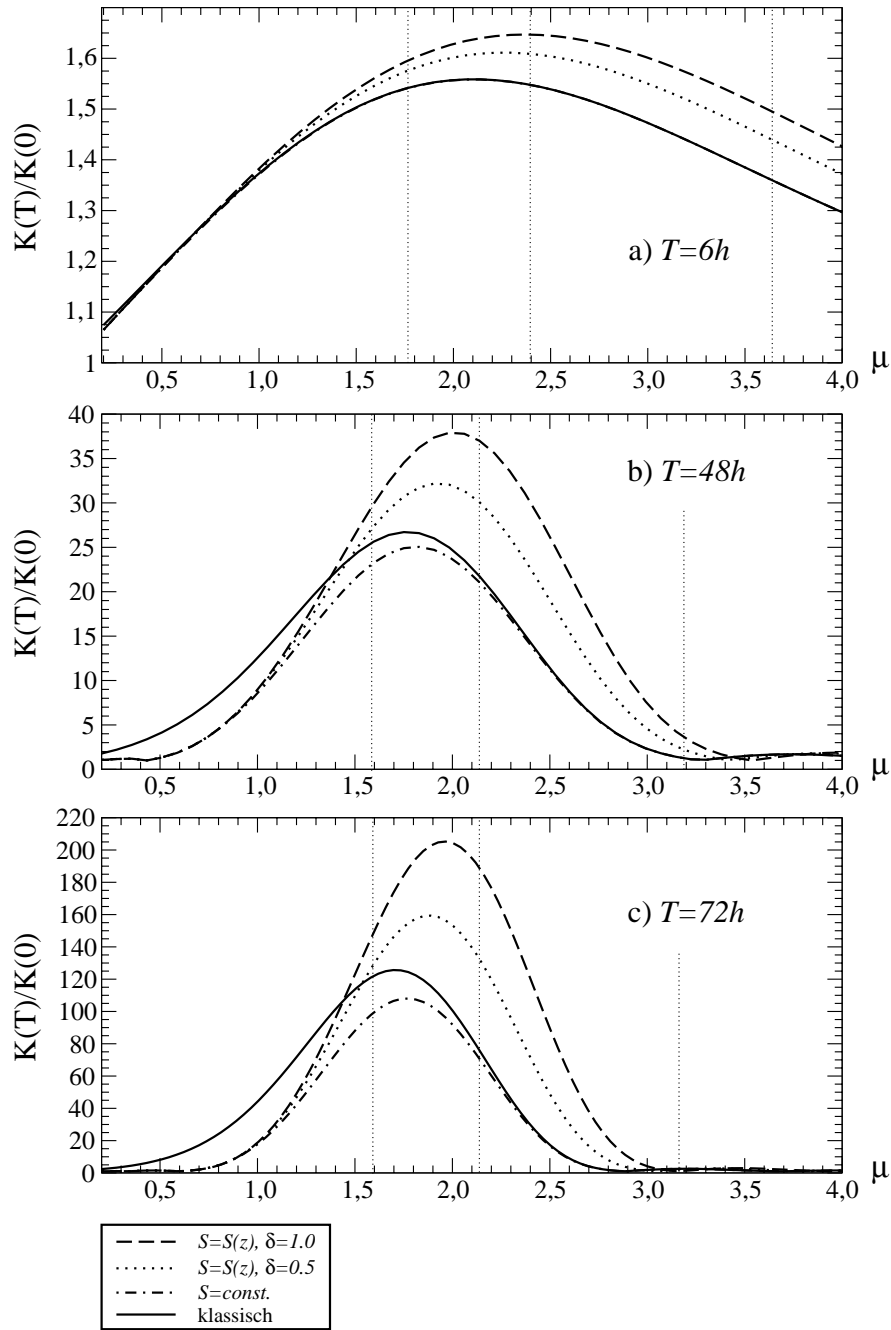


Figure 5: SIA - growth rates for $T = 6h$, $T = 48h$ and $T = 72h$.

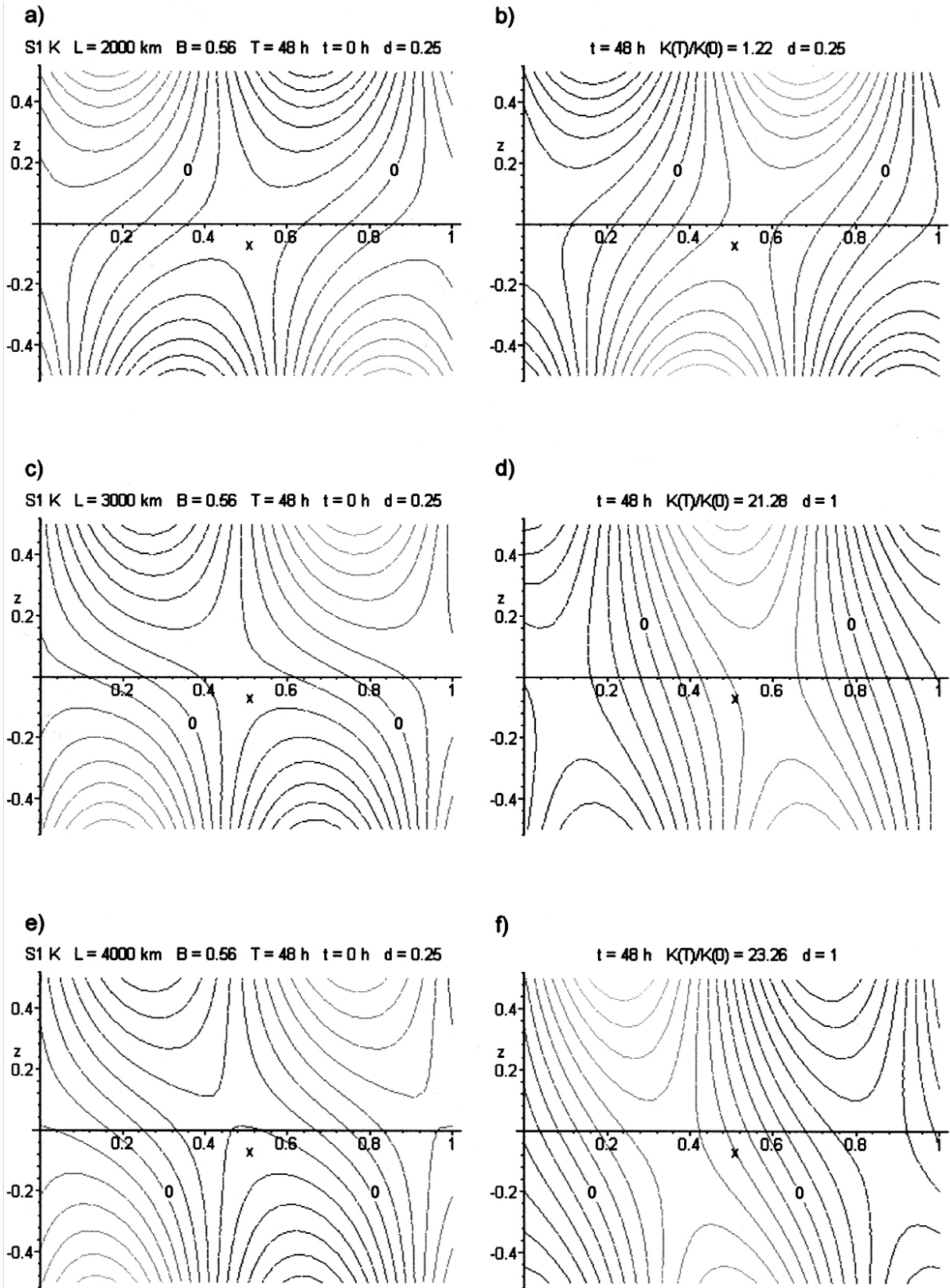


Figure 6: SIA - structures for $T = 48h$ and the wavelength $L = 2000km$, $L = 3000km$ and $L = 4000km$.

for this wavelength is the $S = S(z)$ -model with $\delta = 1.0$ (growth rate = 3.54). This is due to the destabilization in the height with increasing δ , and the fact that the period of oscillation fits very well to optimization interval in this case.

$L = 3000km$ (Figure 6c,d): In contrast to the corresponding S1A-vectors the S1K-vectors show no normalmode-like structure - in agreement with the growth rates discussed before. For $t = 0$ the phase tilts are only slightly smaller than π ($\approx 0.8\pi$ in the classical case), i.e. the minima of the upper wave are nearly above the maxima of the lower wave, what is equivalent to a high contribution of the sinh-basis function to the structure - corresponding to little K and much A . For $t = T = 48h$ the phase tilt is much smaller, what is equivalent to a higher contribution of the cosh-basis function. This corresponds to little A and much K .

$L = 4000km$ (Figure 6e,f): At this wavenumber one sees well that the contribution of the sinh-basis function to the structure for $t = 0$ is smaller as in the case discussed before. Also the contribution of the cosh-basis function for $t = T$ is smaller. So we see that $A - K$ - conversion is not as effective as in the case before. But nevertheless the growth rate is larger. This is due to the fact, that the phase tilt does not decrease as fast as in the case before, and therefore the gain of energy from the basic flow is larger.

At both wavenumbers the saddle point of the structures descends with time - just as the saddle point of the S1A-structures. But in contrast to the S1A-vectors the starting points of the S1K-vectors is not descending with δ .

For longer optimization intervals the discussed characteristics qualitatively remain the same, and become clearer.

4 Conclusions

Qualitative differences to Fischer's results were only found at small wavenumbers below the longwave cutoff. The most remarkable difference beyond the longwave cutoff is the behaviour of the saddle points of the models with $\beta \neq 0$, which descends with time. That means, that the singular vectors of the models with $\beta \neq 0$ grow faster in the upper half than in the lower half of the fluid for the considered mean flows. According to equations (25) and (40) this behaviour is very plausible.

The growth rates in parameter ranges of neutral normalmodes are too small to enable non-linear growth effects in meteorologically relevant times.

We used a new point of view to optimal growth of structures, which only consist of neutral normalmodes. An appropriate understanding of optimal growth in parameter ranges of neutral normalmodes, can be achieved by calculating of two quantities: the maximal growth rate ("envelope"), which can be reached by singular vectors for fixed parameters (we showed that this quantity is time independent) and the minimal time, which is necessary to obtain this growth rate. This time is the half of the oscillation period given by equation (45). A closer inspection of the envelopes yields, that they go to infinity at the cutoffs. So for every given value of a growth rate a wavenumber beyond the shortwave cutoffs or below the longwave cutoffs can be found at which this growth rate can be achieved by superposition of neutral normalmodes. But on the other hand a closer inspection of the time, which is necessary to achieve the growth rate given by the envelope, yields, that it goes to infinity, too. For this reason the cutoffs must be understood not as sharp border of instability but more as a smooth transition to instability.

References

- [1] **Charney, J. G., 1947**, The dynamics of long waves in a baroclinic westerly current, *J. Meteor.*, 4, 135-163.
- [2] **Eady, E. T., 1949**, Long waves and cyclone waves, *Tellus*, 1, 33-52.
- [3] **Farrell, B. F., 1982**, The initial growth of disturbances in a baroclinic flow, *J. Atmos. Sci.*, 39, 1663-1686.
- [4] -, **1984**, Modal and nonmodal baroclinic waves, *J. Atmos. Sci.*, 41, 668-673.
- [5] -, **1988**, Optimal excitation of baroclinic waves, *J. Atmos. Sci.*, 46, 1193-1206.
- [6] **Faulwetter, R., 2000**, Singuläre Vektoren des 2d - Eady - Modells, Diploma Thesis, University of Leipzig, available from the author.
- [7] **Fischer, C., 1998**, Linear amplification and error growth in the 2d - Eady - problem with uniform potential vorticity, *J. Atmos. Sci.*, 55, 3363-3380.
- [8] **Heinrich, H., 1999**, Optimal wachsende Störungen in einem barotropen Modell, Diploma Thesis, University of Leipzig, available from the author.
- [9] **Joly, A., 1995**, The stability of steady fronts and the adjoint method: nonmodal frontal waves, *J. Atmos. Sci.*, 52, 3082-3108.
- [10] -, **and A. J. Thorpe, 1991**, The stability of time dependent flows: An application to fronts in developing baroclinic waves, *J. Atmos. Sci.*, 48, 163-182.
- [11] **Lindzen, R. S. 1994**, The Eady - problem for a basic state with zero PV gradient but $\beta \neq 0$, *J. Atmos. Sci.*, 41, 3221-3226.
- [12] -, **B. F. Farrell and D. Jaqmin, 1982**, Vacillations due to wave interference. Applications to the atmosphere and to annulus experiments, *J. Atmos. Sci.*, 39, 14-23.
- [13] **Pedlosky, J., 1964**, An initial - value problem in the theory of baroclinic instability, *Tellus*, 16, 12-17.
- [14] **Rotunno, R., and M. Fantini, 1989**, Petterssens "type B" cyclogenesis in terms of discrete, neutral Eady modes, *J. Atmos. Sci.*, 46, 3599-3604.

Address of authors:

Robin Faulwetter, Werner Metz
Universität Leipzig, Institut für Meteorologie
Stephanstr. 3
04103 Leipzig

**Wissenschaftliche Mitteilungen aus dem Institut für Meteorologie der
Universität Leipzig**

- Band 1 *A. Raabe, G. Tetzlaff* und *W. Metz* (Edn.):
Meteorologische Arbeiten aus Leipzig I
- Band 2 *R. Devantier*: Wolkenbildungsprozesse über der südwestlichen Ostsee -
Anwendungen eines neuen Wolkenschemas in einem mesoskaligen Modell
- Band 3 *J. Laubach*: Charakterisierung des turbulenten Austausches von Wärme,
Wasserdampf und Kohlendioxid über niedriger Vegetation anhand von
Eddy-Korrelations-Messungen
- Band 4 *A. Raabe* und *J. Heintzenberg* (Edn.):
Meteorologische Arbeiten aus Leipzig II
- Band 5 Wind- und Seegangsatlas für das Gebiet um Darß und Zingst
D. Hinneburg, A. Raabe und *G. Tetzlaff*: Teil I: Windatlas
- Band 6 *W. von Hoyningen-Huene* und *G. Tetzlaff* (Edn.): Sediment and Aerosol
Teil I: Beiträge zur Alfred-Wegener-Konferenz, Leipzig 1997
Teil II: Aktuelle Beiträge aus dem Institut für Meteorologie
- Band 7 *B.-R. Beckmann*: Veränderungen in der Windklimatologie und in der
Häufigkeit von Sturmhochwassern an der Ostseeküste Mecklenburg-
Vorpommerns
- Band 8 *P. Posse*: Bestimmung klimarelevanter Parameter des maritimen Aerosols
unter besonderer Berücksichtigung der Nichtkugelform realer Aerosolteilchen
- Band 9 *A. Raabe, K. Arnold* und *J. Heintzenberg* (Edn.):
Meteorologische Arbeiten aus Leipzig III
- Band 10 Wind- und Seegangsatlas für das Gebiet um Darß und Zingst, Teil II:
D. Hinneburg, A. Raabe und *G. Tetzlaff*: Vergleich Windatlas –
Beobachtungsdaten; *M. Börngen, H.-J. Schönfeldt, F. Riechmann, G. Panin*
und *G. Tetzlaff*: Seegangsatlas; *M. Stephan* und *H.-J. Schönfeldt*:
Sedimenttransportatlas
- Band 11 *J. Reißmann*: Der Einfluß langwelliger Strahlungsprozesse auf das
bodennahe Temperaturprofil
- Band 12 *A. Raabe, K. Arnold* und *J. Heintzenberg* (Edn.):
Meteorologische Arbeiten aus Leipzig IV
- Band 13 *U. Müller, W. Kuttler* und *G. Tetzlaff* (Edn.):
Workshop Stadtklima 17. / 18. Februar 1999 in Leipzig

- Band 14 *R. Surkow*: Optimierung der Leistungsverfügbarkeit von Windenergie durch ihre Integration in Wind-Biogas-Hybridanlagen
- Band 15 *N. Mölders*: Einfache und akkumulierte Landnutzungsänderungen und ihre Auswirkungen auf Evapotranspiration, Wolken- und Niederschlagsbildung
- Band 16 *G. Tetzlaff* und *U. Grünewald* (Edn.):
2. Tagung des Fachausschusses Hydrometeorologie 15./16. 11. 1999 in Leipzig
- Band 17 *A. Raabe* und *K. Arnold* (Edn.):
Meteorologische Arbeiten aus Leipzig V
- Band 18 *K. Arnold*: Ein experimentelles Verfahren zur Akustischen Tomographie im Bereich der atmosphärischen Grenzschicht
- Band 19 *A. Ziemann*: Eine theoretische Studie zur akustischen Tomographie in der atmosphärischen Grenzschicht
- Band 20 *Ch. Jacobi*: Midlatitude mesopause region dynamics and its coupling with lower and middle atmospheric processes
- Band 21 *M. Klingspohn*: Interdekadische Klimavariabilität über dem Nordatlantik – Statistische Analysen und Modellstudien –
- Band 22 *A. Raabe* und *K. Arnold* (Edn.):
Meteorologische Arbeiten aus Leipzig VI