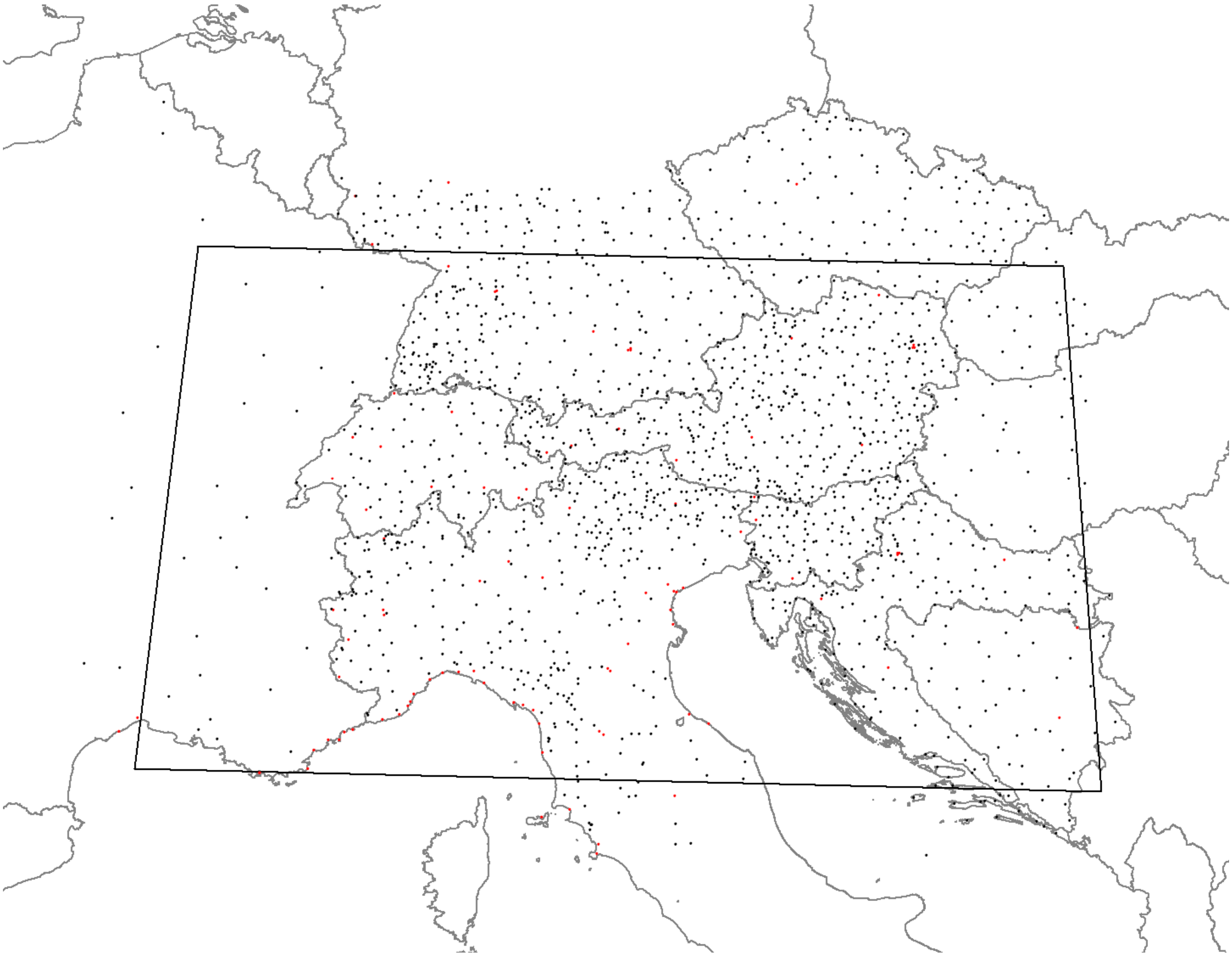
ECSN – HRT/GAR

High Resolution Temperature Climatology in Complex Terrain – demonstrated in the test area Greater Alpine Region GAR

Annex 1 to Final Report

DATA PROCESSING
FOR A HIGH RESOLUTION ALPINE TEMPERATURE CLIMATOLOGY
(ECSN/HRT-GAR)



Vienna, June 2008

DATA PROCESSING
FOR A HIGH RESOLUTION ALPINE TEMPERATURE CLIMATOLOGY
(ECSN/HRT-GAR)

1. The original data (data sources, providers…)

The initial phase of the Alpine temperature mapping activity was characterised by an intensive search for all existing regular temperature data for the GAR (Greater Alpine Region). The leading principle was to achieve the highest possible spatial density of monthly temperature means for the period 1961-90. High data density is fundamental due to the strong vertical structure of the Alps which causes sub-grid-effects anyway easier to be modelled when network density allows for representative and significant regression analysis also in sub-regions like coasts, cities, inner-alpine basins and others.

The 1961-90-period was chosen because weather services had already invested work into this official WMO-CLINO period. A more recent period (1971-2000 for example) would have caused problems in respect to the inclusion of the 1990s (especially in Italy where a fundamental re-organisation of the existing networks took place in the 1990s). A longer period (e.g. 20th century) would have caused other severe problems in terms of the gaps during the 2 world wars, sparser networks in general and organisational and political breaks. The 1961-90 CLINO-sample is not necessarily the ultimate goal. The 1971-2000 climatology can produced better by working on differences between 1961-1990 and 1971-2000 by including all ECSN/HRT-GAR findings. It is planned to produce any other climatology in a further step (see chapter 6. Outlook). Extensions much further back (late 18th century) can be performed by merging ECSN/HRT-GAR climatology with the HISTALP long-term dataset (Auer et al., 2007) using a similar approach as described for precipitation by Efthymiadis et al., 2005.

To achieve a maximum high spatial density, another precondition was handled less restrictive – full 30-years temporal coverage without gaps was not defined as binding condition. We accepted also datasets with gaps and such which covered earlier or later samples than 1961-90 if highly correlated neighbouring datasets allowed for filling the gaps or adjusting to the reference period - for both purposes a stability of the mean differences between highly correlated series was assumed.

The climatologically interesting but politically and administratively highly scattered "Greater Alpine Region" afforded a time consuming period for data detection, organisation and also digitising. The data collecting finally resulted in about 1800 single station datasets of monthly mean temperatures (t_m) or mean daily extremes (t_x and t_n) from various providers in the region. Details are shown in Table 1

Table 1: Overview of data and its sources

Country	data source
Austria	ZAMG climate data base
Austria	Yearbooks of Hydrographical Service of Austria
Bosnia & Herzegovina	METEOBiH / Zelko Majstorovic
Bosnia & Herzegovina	Yugoslavian Yearbooks
Croatia	DHMZ / Ksenija Zaninovic
Czech Republic	CHMI / Vit Kveton
Germany	DWD / Gerhard Müller - Westermeier
France	Meteo France / Jean-Marc Moisselin
Hungary	OMSZ / Zita Bihari
Italy	Marco Carrer / University of Padua
Italy	Maurizio Maugeri / University of Milano
Italy	Società Meteorologica Italiana
Italy	Paola Nola; Renzo Motta / Universitiy of Pavia, University of Torino
Italy	Giancarlo Rossi
Italy	University of Torino / CD
Slovakia	SHMU / Oliver Bochnicek
Slovenia	ARSO / Mojca Dolinar
Switzerland	Meteo Swiss / Michael Begert

2. Station locations (coordinate corrections, urban, rural and other specific environments)

The original number of collected single station datasets was approximately 1800 but not all of them finally met the requirements in terms of completeness and data-quality. In a first step duplicates referring to the same site had to be eliminated (which was not clear in each case before a closer look at site-names and coordinates, queries at the data-providers,...). Some of the duplicates were different datasets under the same name and needed a further specification (typically cities with more than one measuring site).

A major problem appeared and had to be solved which is typical for any project which targets at high-resolution spatial climate analysis - the problem of coordinate errors. In former times, the exact geographical position of a measuring site was regarded essential more for the altitude, less for the longitude and latitude parameters. The now existing highly resolved elevation and land-use grids (we planned to use a resolution of at least 1km plus an algorithm to deal with sub-grid elevation variability within) now afford a much more precise anchoring of the measuring sites into the elevation and land-use models to fully exploit their now given wider potential through GIS-based analysis. Examples are the inclusion of land-use effects on temperature (urban-rural, glaciers, lakes, etc), the influence of the local land-forms (basins, slopes, summits,) and a visual residual analysis of the different steps of the modelled t_m-fields versus the point information of the single sites. All those (and a number of other applications) afford precise information on longitude, latitude and altitude of stations.

It was an absolute necessity for further analysis to check each single site location because the majority of coordinates had intolerable location errors for the planned application. The errors were nearly exclusively referring to the longitude and latitude, altitude was much more accurate. The vast majority of errors originated from the fact that many data providers only used geographical degrees and minutes and no subdivision into seconds. This alone shifts some summit stations into nearby valleys, coastal sites into the sea, urban sites into rural surrounding and other intolerable facts.

Of course, it was not possible to visit the real sites and check their location. It was a combined approach of using the station history files and yearbooks, contacting the data providers, and scanning printed maps and digital sources like Encarta, Google-Earth and also internet requests for single sites (e.g. mountain resorts, summits etc...).

3. Adjustment to common 1961-90 reference period (gap-filling and adjustments from other than 1961-90 samples)

Taking into account the existing long-term temperature variability in the region (Fig.1) it is essential to refer all data to a common period.

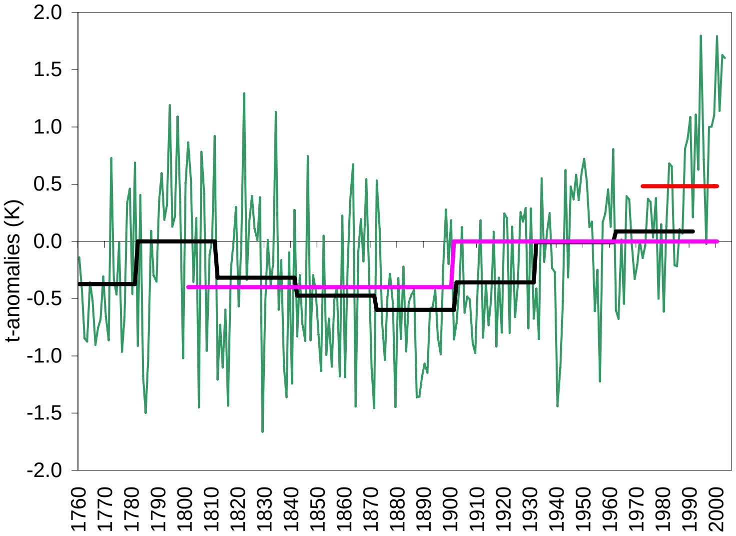


Figure1. Annual mean temperatures in the GAR 1760-2003 (anomalies from 1901-2000), green: single years, black: 30-year-CLINO-means, pink: 19th and 20th century means, red: 1971-2000

As already mentioned, one of the reasons to choose the 1961-90 reference was the incomparably greater number of complete datasets compared to others. For France, Switzerland, Germany, Austria, Czech Republic, Slovak Republic, Hungary, Slovenia and Croatia all monthly averages were already either directly calculated from complete datasets or adjusted to the common 1961-90 reference period. For Italy and Bosnia & Herzegovina some of the sampled datasets had either gaps or did not cover the entire 1961-90 period. Some of them were much longer (in single cases reaching back into the 1920s) but ended before 1990. Apart from the respective HISTALP sites (where gap closing had been done already, Auer et al., 2007) the vast majority of incomplete datasets could be gap-filled and/or adjusted using highly correlated comparative station datasets and assuming constancy of the inter-station t_m-differences.

Finally a network of about 1726 stations (Fig.2) passed the location-correction and the gap-filling and sample adjustment procedures.

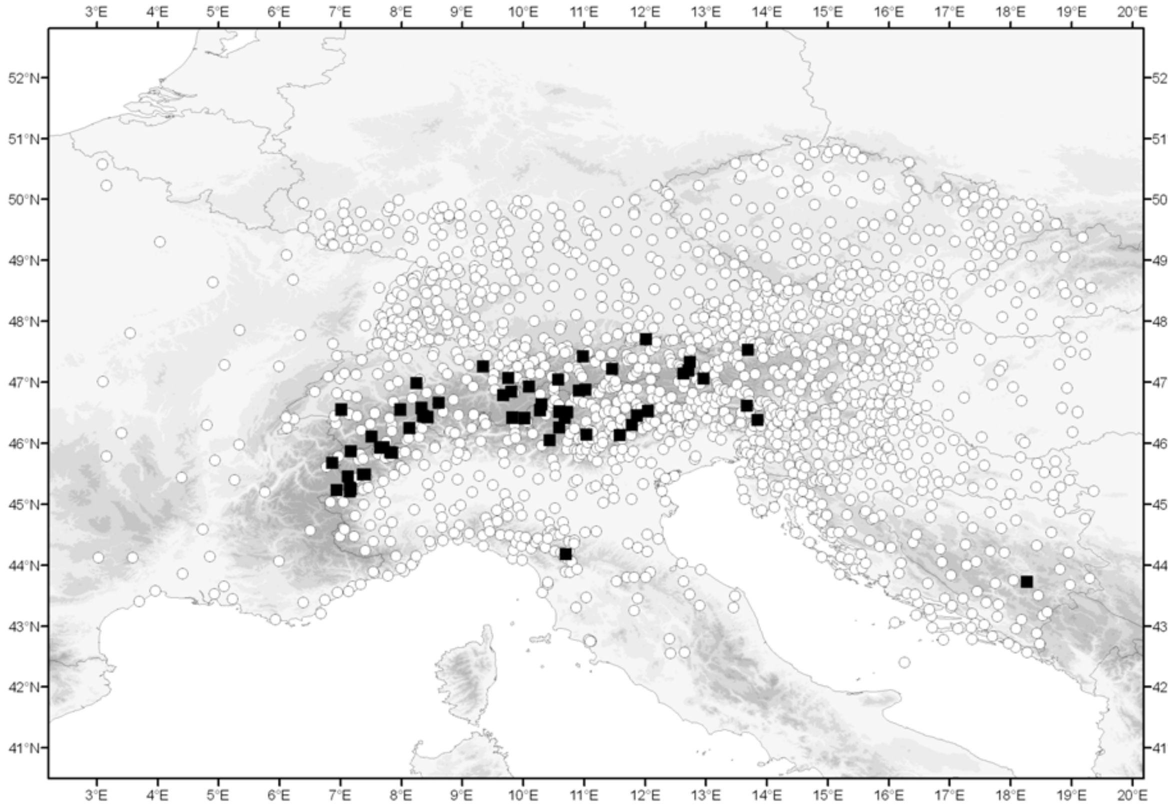


Figure 2: The ECSN HRT-GAR network, black squares indicate mountain stations

4. Adjustment to common means-calculation (regionalisation according to DTR – estimating mean daily courses for sub-regions or single sites – estimation and elimination of means calculation biases)

As indicated in Table 2 the observation times and therefore calculation algorithms were not same for provided data. They had to be adjusted before being analysed.

Table 2: Used algorithms for the calculation of monthly mean temperature in the GAR during 1961-1990 (status of provided data)

(7+14+21+21)/4	Bosnia&Herzegovina	MLT
	Czech Republic	MLT
	Germany 1961-1986	MLT
	Croatia	MLT
	Hungary	MLT
	Slovenia	MLT
	Slovakia	MLT
(7.30 + 14.30 + 2*21.30)/4	Germany 1987-1990	CET
(tx+tn)/2	France	
	Italy	
pre-adjusted to TRMs	Austria	
	Switzerland	

About half of the provided data were based on the Kämtz-Formula (Kämtz, 1860) $t_m = (t_7+t_{14}+2*t_{21})/4$, observing times in MLT

a quarter as “true means” (TRMs), either based on hourly values or pre-adjusted to TRMs

a quarter based on the mean daily extremes $t_m = (t_x+t_n)/2$

Studies for the territory of Austria (e.g. Auer et al., 2001) based on a 10-years 50-stations-dataset (1986-1995) of hourly temperature measurements (henceforth called AT-T24-50) shows that the differences between the 2 estimates and the TRMs are not negligible. Fig.3 gives respective examples for the biases of the Kämtz-formula and the t_x,t_n–estimate in some sub-regions of Austria.

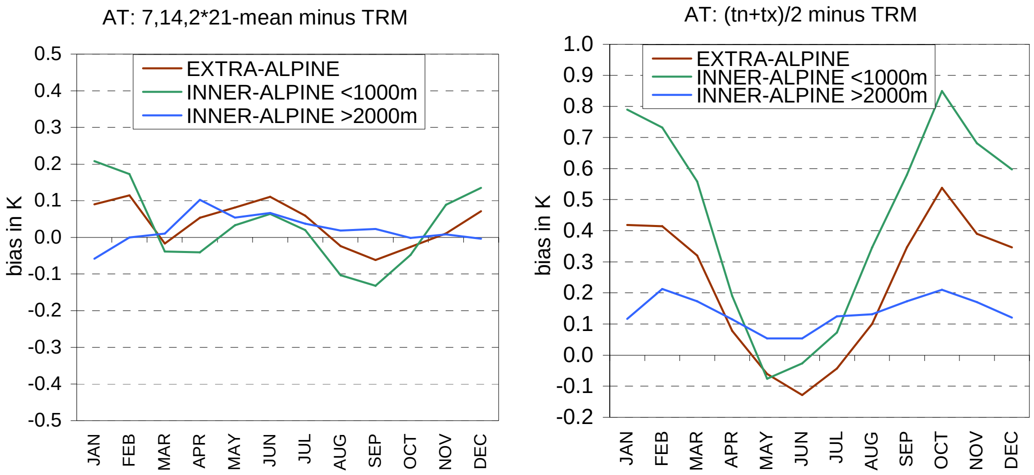


Figure 3: Biases of the two estimates for temperature means calculation in respect to the true mean (TRM) derived from the AT-T24-50 dataset

It is evident that the Kämtz-formula is the better estimate for the true means and its maximum 0.2K deviations from TRM can be tolerated for climate mapping compared to other uncertainties which can be expected to produce a remaining mean residual of point measurements vs. HR-fields of approximately 0.5K (Böhm and Potzmann, 1999). The estimate based on $(t_x+t_{n)}/2$ however exceeds the

0.5K-threshold in some months and must be regarded as too strongly biased to merge it with the other two groups of data present in the dataset. Table 2 indicates another potential bias in means estimation, the use of CET or MLT for climate observations and measurements. We hope to have thoroughly reconstructed the official respective regulations in the different networks. One of the regulations followed everywhere is the (physically meaningful) neglecting of any daylight saving regulations. But there might be also certain discrepancies between the official regulation and the everyday routine at single stations. Anyway the example of Fig.4 of a (fictitious) one hour shift of observing times in summer (which would be the "worst case" for any biases in respect to incorrect use of time regulations) demonstrates that such biases are also very small in respect to an application for climate mapping.

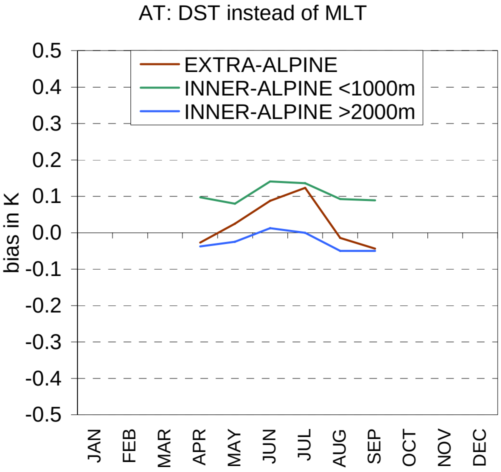


Figure 4: Potential bias of a Kämtz- t_m-estimate if daylight saving time (DST) regulation was not neglected

The final decision on the optimal means-estimate for our application was to use the Kämtz-formula. This allowed about half of the data to be included without change, to apply minor adjustments to those with TRMs and to concentrate on finding feasible solutions for the stronger adjustments necessary for French and Italian stations using (tx+tn)/2.

In detail:

Austria: A simple and direct inclusion of means from ZAMG' and HZB's climate databases was not possible due to the discrepancies and breaks of observing times in the CLINO period All datasets had been gap-filled, adjusted to the 1961-1990-period and to TRM already (Auer et al., 2001b). In addition to that, the described AT-T24-50 dataset of mean daily temperature courses (DTC) in 6 Austrian sub-regions plus a DTR-analysis (mean daily temperature range) is used as the basis to derive respective corrections for the other subsets with biased means. Fig.5 shows typical examples of such mean daily courses for four of the detected six homogeneous sub-regions (2 extra-alpine ones and an inner-alpine one subdivided into in 4 altitude bands). Together with an analysis of DTRs (resulting in a best estimate of a factor of 0.9 between the DTR as the difference of really measured daily extremes and the DTR as the difference of the highest and lowest hourly temperature) these data were used to re-adjust the Austrian temperature means to the Kämtz-formula.

The Austrian DTCs were also taken as the basis to produce DTC-estimates for Switzerland, Italy and France in order to produce Kämtz- t_m–estimates also there (details on the following).

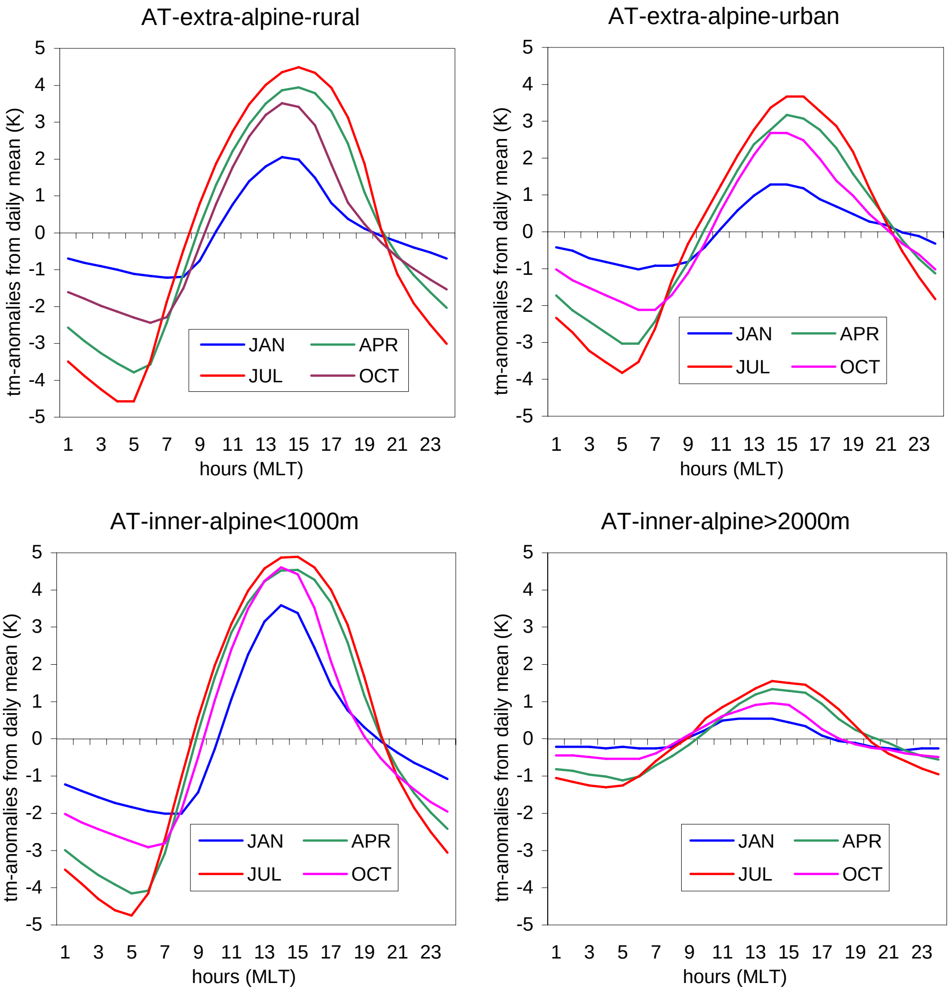


Figure 5: Mean daily temperature-courses (DTCs) for January, April, July and October in 4 Austrian sub-regions (sample: AT-T24-50 dataset)

Switzerland: MeteoSwiss provided a 91-station dataset of 1961-1990 monthly t_m–, t_x– and t_n–means, t_m either directly measured (ANETZ-sites) or pre-adjusted (conventional sites) to TRMs. Therefore a relatively simple procedure could be applied to produce first the respective DTRs and DTCs and in a second step derive from them the adjustments from TRMs to Kämtz-means. For most of the sites (especially low elevation sites) Swiss DTRs (and consequently DTCs) are similar to those in Austria, only at high elevations and in the Ticino daily temperature amplitudes are higher than in the Austrian mountains. Fig.6a shows the respective sub-regional mean DTRs for Switzerland and Austria.

The basic principle for step 2 was the assumed similarity of the relative shape of the DTCs in both countries (later also for the Italian and the French adjustments), the only differences to be caused by DTC-amplitudes (which were estimated by the described multiplication of the real DTRs by 0.9). This vertical shrinking (if DTR is smaller) or stretching (if DTR is wider) of the DTCs is a simple but sufficiently accurate method to produce the necessary DTRs for regions in which only DTRs are measured (like

Italy and France). For Switzerland it was applied too to re-adjust the true means (TRMs) to Kämtz-means and verified versus the existing TRMs from Swiss ANETZ stations. The results were positive (Michael Begert, personal communication) and let the method seem appropriate for application in Italy and France as well. Fig.6b shows the derived adjustments for Swiss sub-regions. They are only slightly higher than those for respective Austrian sub-regions (compare Fig. 3. left) but have the same bimodal shape.

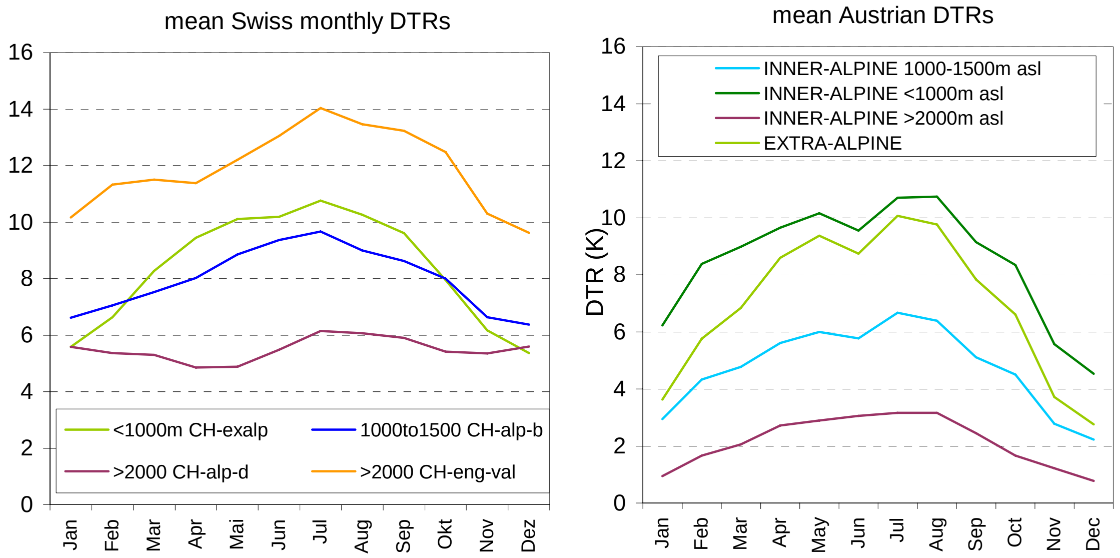


Figure 6a: Mean monthly DTRs for some Swiss (left) and some Austrian (right) sub-regions: Shown are sub-regions 3 (CH-exalp), 2 altitude bands of sub-region 4 and sub-region 5 for Switzerland, sub-region 2 (extra-alpine) and 3 altitude bands of sub-region 1 for Austria, (spatial distribution of DTR-sub-regions is shown in Fig.7)

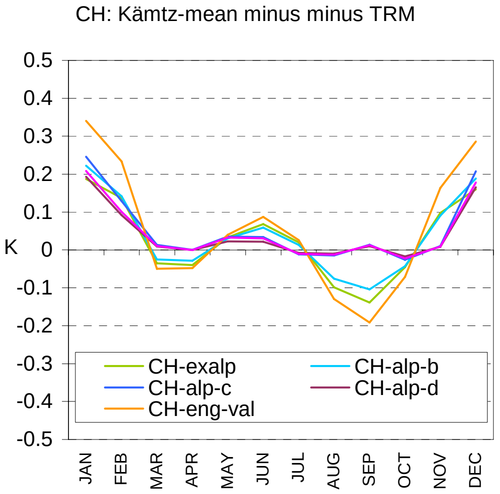
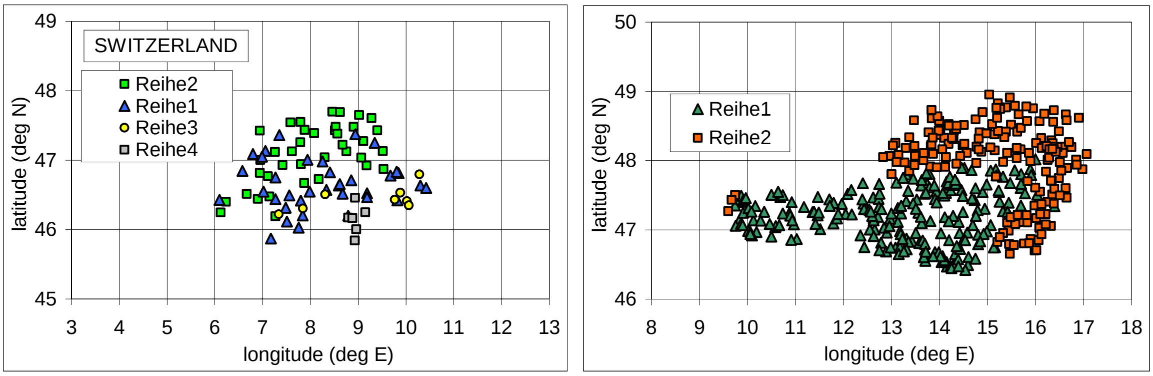
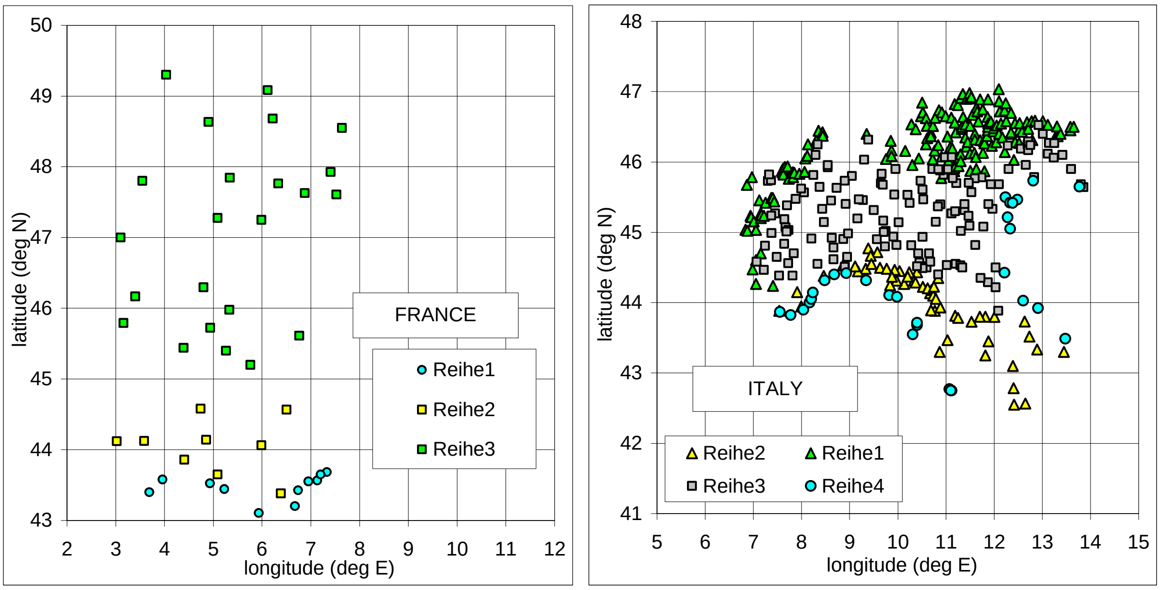


Figure 6b: Biases of the Swiss means calculated after the Kämtz formula compared to true means (TRMs)



left: Switzerland: 1: Alps (4 alt.bands), 2: ex-alp., 3: Engadin-Valais, 4: Ticino
right: Austria: 1: Alps (4 alt.bands), 2: ex-alp



left: France: 1: coastal, 2: Mediterr.-inland, 3: NE-France
right: Italy: 1: Appennino-Toscana (4 alt.bands), 2: Ital.Alps (4 alt.bands),
3: N-Italian plains (inland), 4: coastal

Figure 7: Regionalisation of Swiss, Austrian, French and Italian sites into homogeneous sub-regions in terms of DTR and DTC

Italy: In Italy, a subset of 191 monthly mean t_x- and t_n-stations (all for-, or adjusted to 1961-90) could be used to apply a procedure similar to the one for the Swiss subset. The difference was that not each site had also the tx- and tn-values available and that the adjustments had to be applied from $(t_x+t_{n)}/2$-means to the Kämtz-algorithm. The former limitation afforded a regionalisation to produce sub-regional mean DTRs and DTCs for those sites with t_m-values alone. The part of Italy present in the GAR had to be subdivided into 4 main DTR-regions, two of them (in the Alps and in the Appennino-Toscana region further subdivided into different altitude bands.

Most of the Italian sub-regions show higher DTRs than the Swiss and Austrian ones from north of the Alps, specifically higher in the Toscana-Appennino region. Only the coastal sites have reduced DTRs due to reduced nightly cooling in the moist maritime air together with the land-sea-wind system which also tends to reduce DTR. Apart from the DTRreduction with increasing altitude (steeper in the Apennines), also the Po-plain (subregion 8) has reduced DTR compared to the low-elevation parts of the Alps and the Apennines, obviously a result of dust and fog in this flat region.

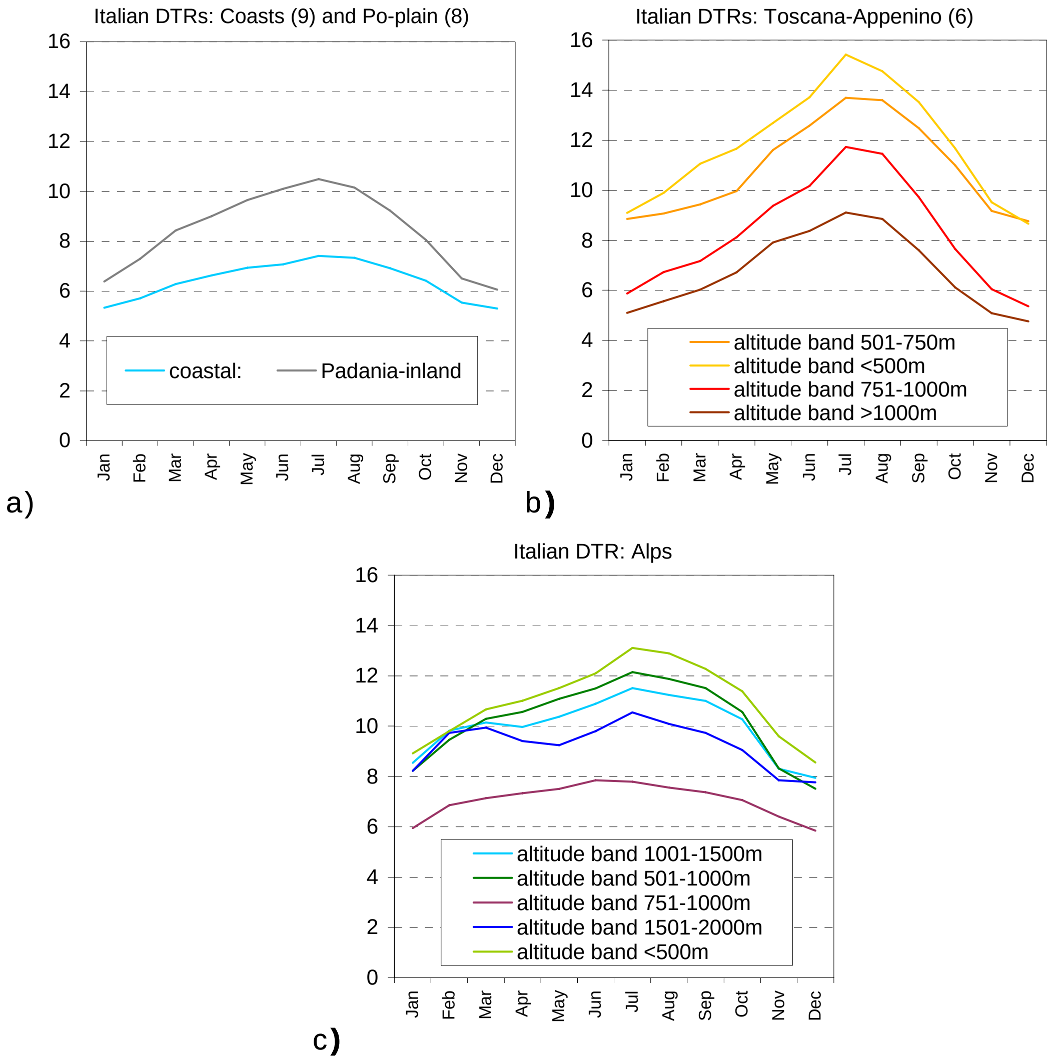
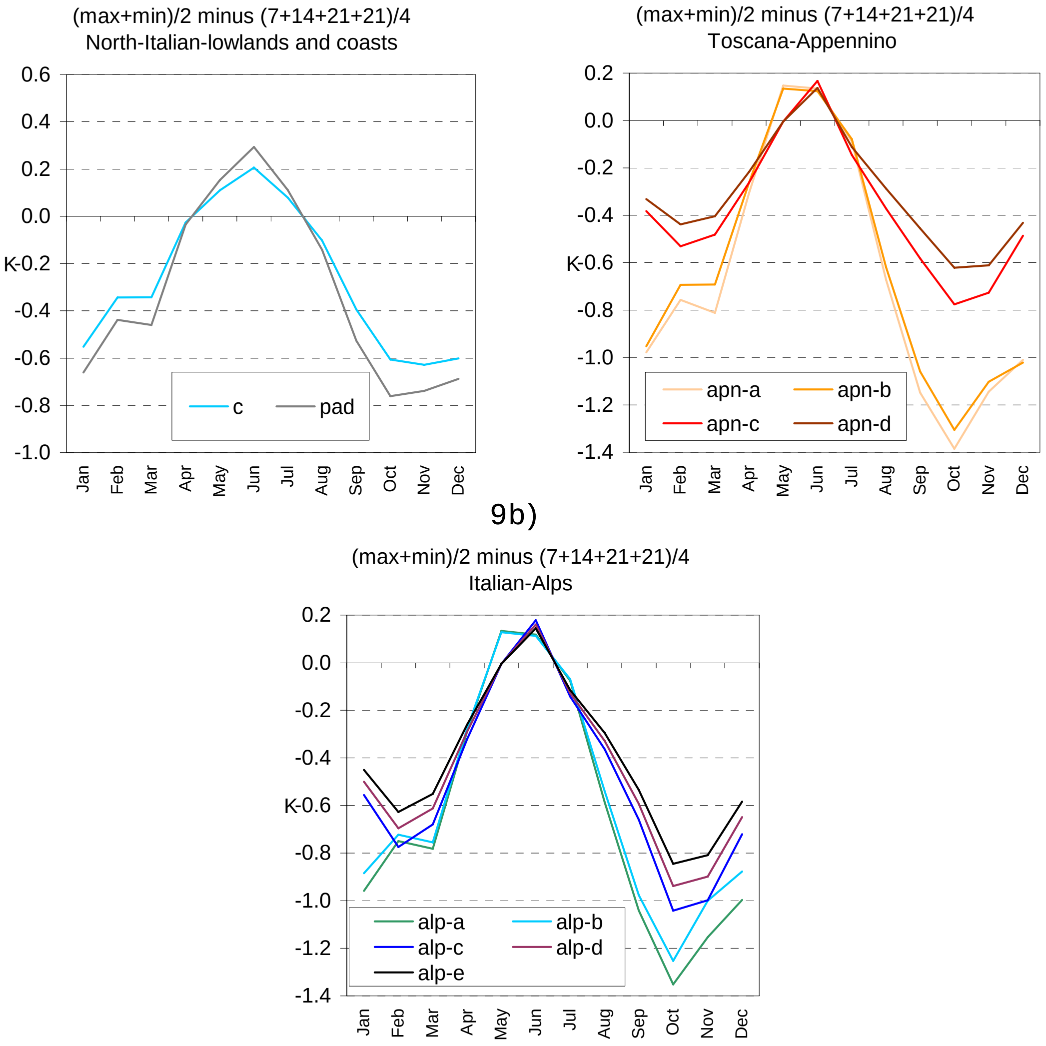


Figure 8:Mean monthly DTRs for the Italian sub-regions : shown are sub-regions 8 and 9 (Fig.6a), sub-regions 6 (Fig.6b) and 7 (Figure 6c), (spatial distribution of DTR-sub-regions shown in Fig.7)

Taking into account the higher DTR-values in Italy and the general deficiencies of the $(t_x+t_n)/2$-approach, it is not surprising that also most of the necessary adjustments to be applied on the Italian $(t_x+t_n)/2$-means to produce the Kämtz-means are larger than those in Austria and Switzerland. Fig. 9 tells that the adjustments span a range from +0.3K to -0.7K in the lowlands and at the coasts, an even wider one in the Italian Alps and the Toscana-Appennino region from +0.2 to -1.4K. The strongest deviations of the $(t_x+t_n)/2$-means from the Krämtz-means are given for the cold season, specifically for October. The vertical structure of the adjustments is steeper in the Apennines than in the Alps, possibly due to the lower altitudes of the Apennines (which cause a higher share on summit sites with reduced DTC at lower altitudes already compared to the Alps – a similar effect to the respective Austrian-Swiss-differences).



9a) 9b)

9c)

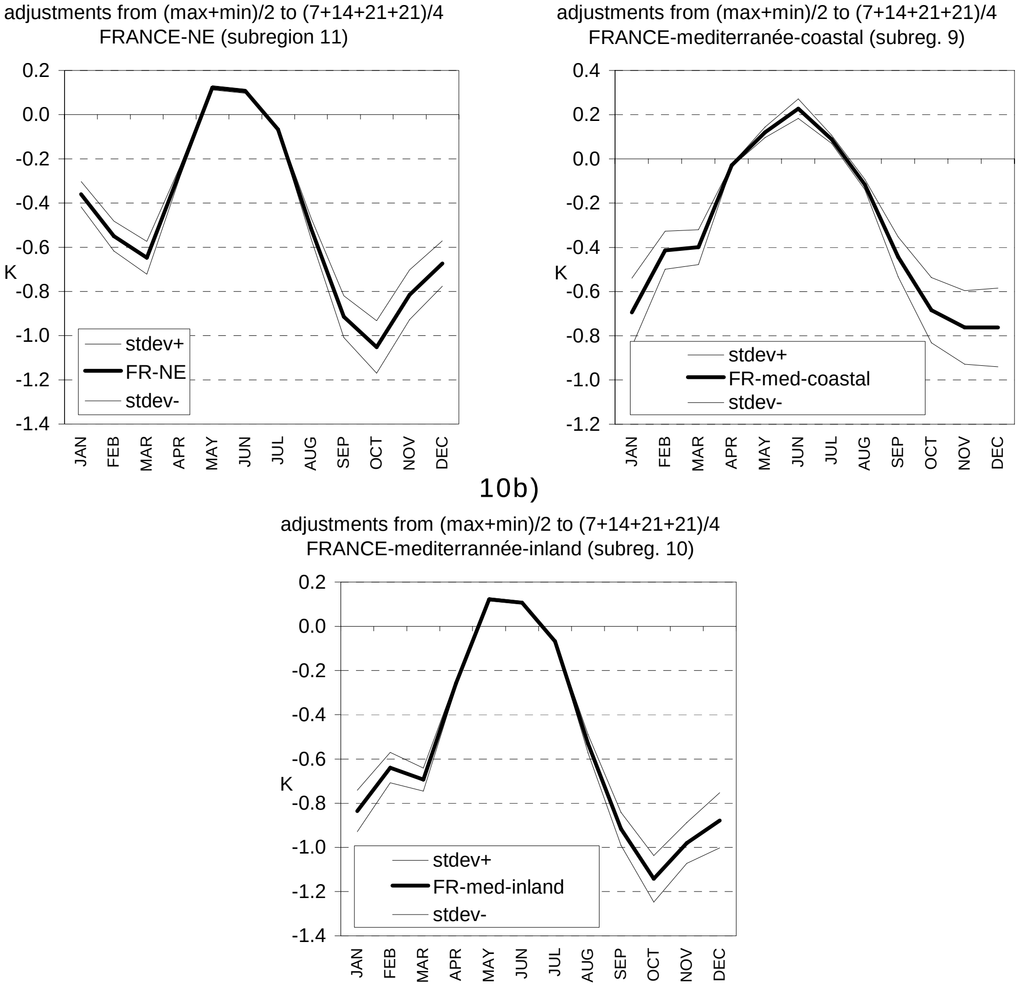
Figure 9: Adjustments to be applied on the Italian (t_x+t_n)/2-means to estimate Kämtz-means. 9a shows sub-regions 8 (pad) and 9 (c), 9b shows sub-region 6 in four 500m-altitude bands from a (<500m) to d (>1500m), 9c shows sub-region 7 in five 500m-altitude bands from a (<500m) to e (>2000m). (Spatial distribution of sub-regions shown in Fig.7)

Compared to the adjustments for the Swiss and Austrian data north of the Alps, the Mediterranean climate south of the Alps – together with the less accurate (t_x+t_n)/2-means estimate – made it much more necessary here to adjust the provided means to achieve comparable data to the majority of the other GAR-subsets.

France: MétéoFrance provided 113 t_x- and t_n-station-datasets which had been quality improved and adjusted to the 1961-90 reference period during the "Le Climat de la France"-project (MétéoFrance, 1999). 45 of them (those east of 3 deg E) were included into the GAR-dataset. A regionalisation of Eastern France in terms of DTR resulted in a subdivision into 3 parts (visible also in Fig.7):

- a large sub-region (11) covering NE-France (approximately north of 45 deg N and east of 3 deg E)
- a Mediterranean sub-region (10) south of 45 deg N but excluding the coasts
- Mediterranean coasts of France east of 3 deg E (9)

The high similarity of the French sub-regions 11/10 to the Swiss/Italian sub-regions 3/6 advised to use the respective adjustments derived from the Swiss and Italian regional mean DTRs and DTCs (based on much denser networks) as a basis to calculate the respective single station adjustments necessary for the two French non-coastal sub-regions. The French coastal sites had been included in the respective analysis of sub-region 9 already and the respective DTRs and DTCs could be used without change anyway.



10a) 10b)

10c)

Figure 10: Adjustments to be applied on the French $(t_x+t_n)/2$-means to estimate Kämtz-means: regional means (bold) plus the 1-standard deviation range (thin lines). 10a shows sub-region 11 (French-NE), 10b shows sub-region 9 (French-Mediterranean coasts, 10c shows sub-region 10(French Mediterranean inland). (Spatial distribution of sub-regions shown in Fig.7)

Fig. 10 shows the resulting adjustments for the three French sub-regions in a slightly different style than those for Austria, Switzerland and Italy (where only the means were included). It shall underline that the adjustments were calculated for each single site. The relatively small 1-standard deviation range (the thin lines) confirms the alternative to use regional mean adjustments as well without severe quality reductions. The general features of the adjustments are similar to those: largest for the Mediterranean inland sites with values of more than 1K in some months, slightly reduced for coastal sites and reduced to less than 1K for each month for the non Mediterranean NE of France.

5. Description of the final version of the alp-map-1961-90 dataset (station map, station list, provider statistics, two first raw regionalising alternatives: (tm(long), tm (lat) and tm(alt) vs. coarse sub-regions

After having corrected, completed and adjusted the original data to a common reference period and to a common means estimate, a final dataset of 1726 single sites monthly temperature means has been made available for the high-resolution gridding. Experience from national or regional gridding activities suggested an initial elimination of urban influenced sites and such with temporarily not yet clearly understood stronger deviations from preliminary t_m–altitude models. 37 urban suspected too warm stations were detected and 25 with cold or warm biases. The former have been used later to produce an additive urban heat-islands-field based on urban-rural temperature differences and on a HR-land-use model. The latter have been dedicated to closer examination: some may be completely eliminated due to remaining error suspicions, some may indicate regional special regions like for example the high Engadin or some other high elevated flat basins or wide valleys with remarkable negative temperature deviations mainly in winter. The last Figures shown below shall provide a first outline of the given spatial temperature variability in the GAR which will have to be reduced to produce the aspired modelled monthly temperature fields. It will be vertical, latitudinal and longitudinal gradients as well as special features like land-sea effects and others which will have to be handled.

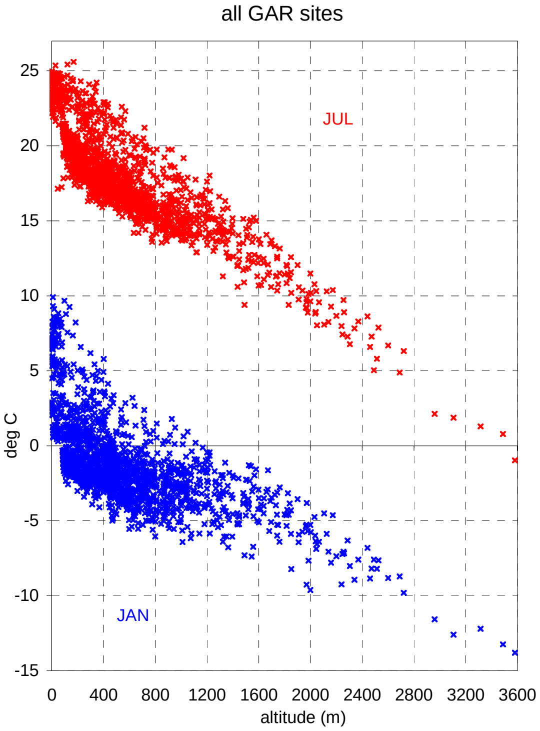


Figure 11: Scatter diagram of January and July mean temperatures for the entire GAR versus altitude (sample with elimination of urban and bias-suspected stations)

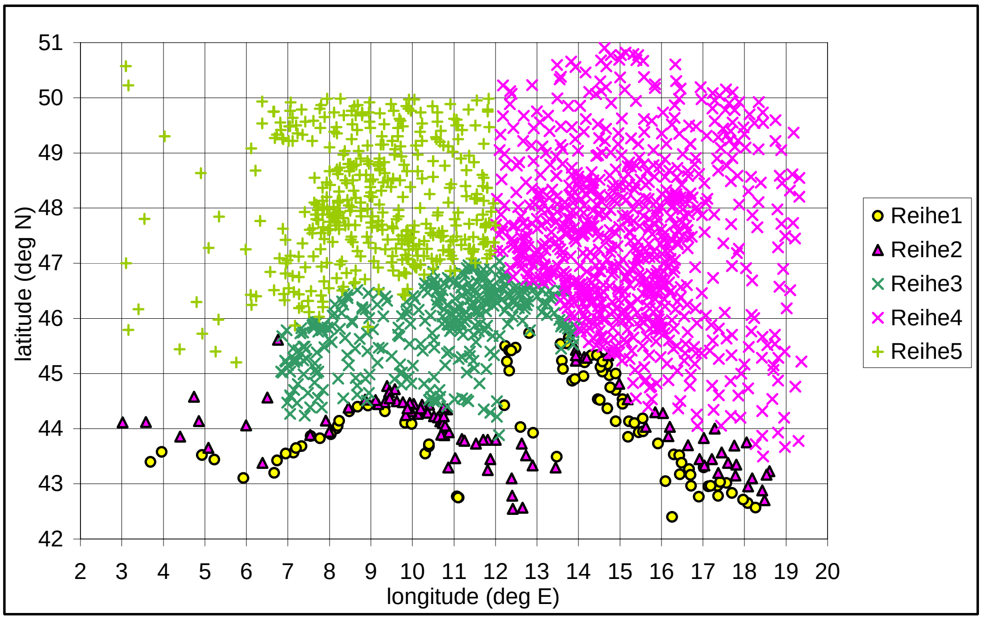


Figure 12. Coarse resolution sub-regions of the GAR: 1: C (coastal), 2: MM (Mediterranean mountains), 3: MMC (N-Italian Inland), 4: NE (northeast-continental.), 5: NW (northwest-Atlantic)

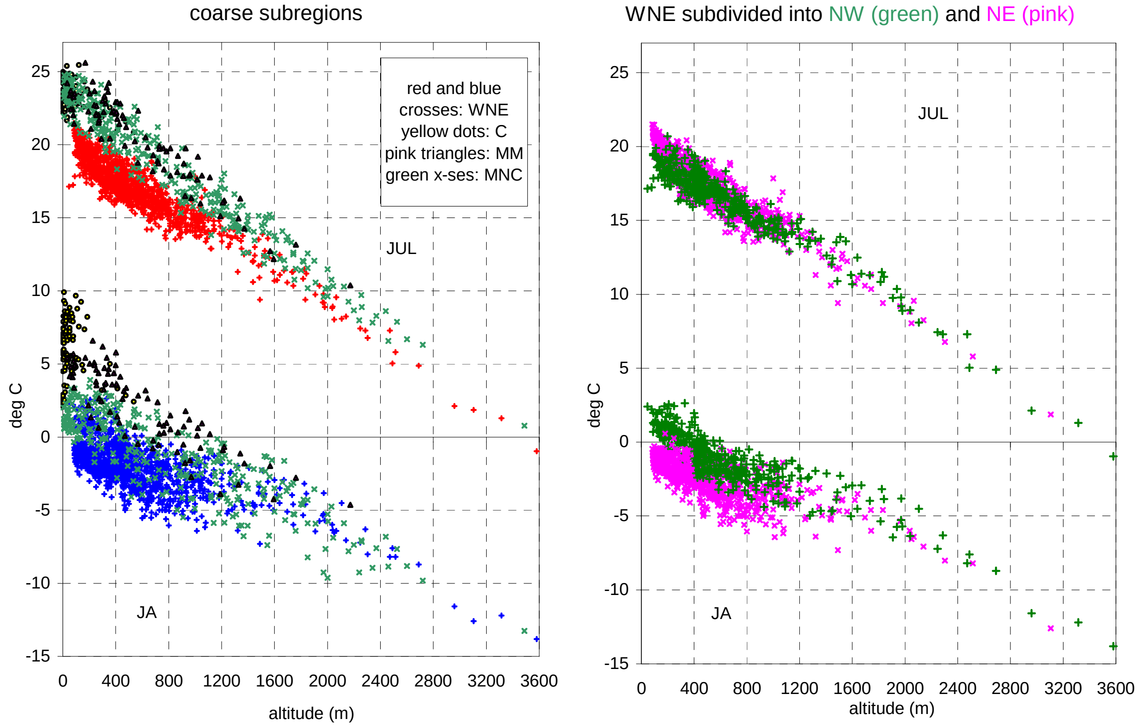


Fig.13: Scatter diagram of January and July mean temperatures for the coarse resolution sub-regions of the GAR versus altitude, left: 4 leading sub-regions, right: sub-region WNE further subdivided into NW and NE. (sample after elimination of urban and bias-suspected stations, spatial distribution of the sub-regions shown in Figure 12)

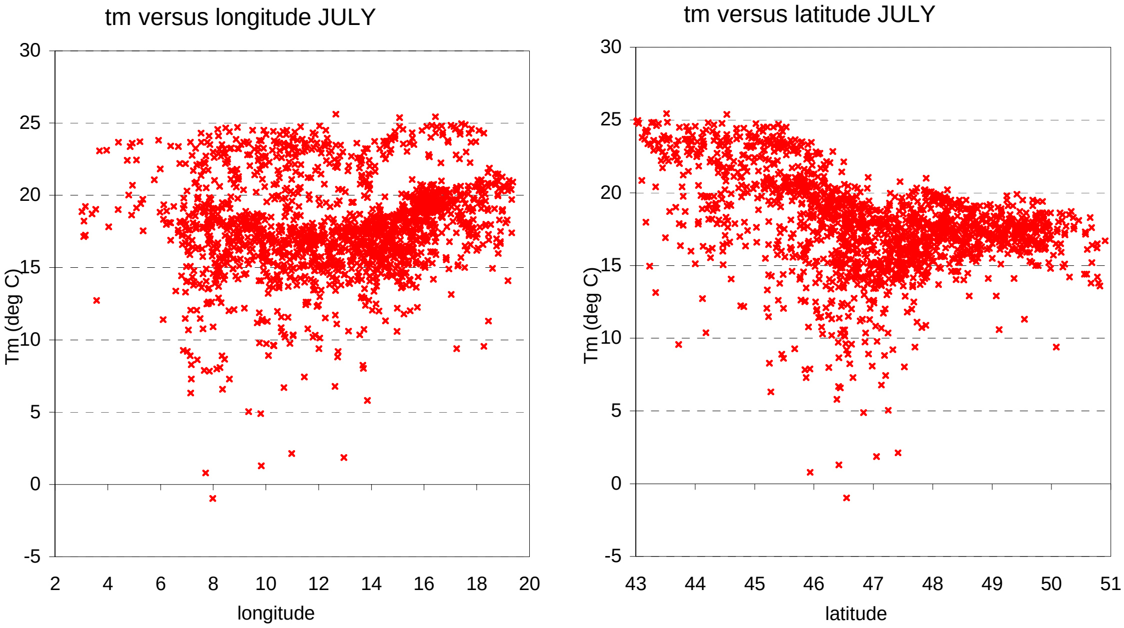


Figure 14: Scatter diagram of July mean temperatures for the entire GAR versus longitude (left) and latitude (right), (sample after elimination of urban and bias-suspected stations)

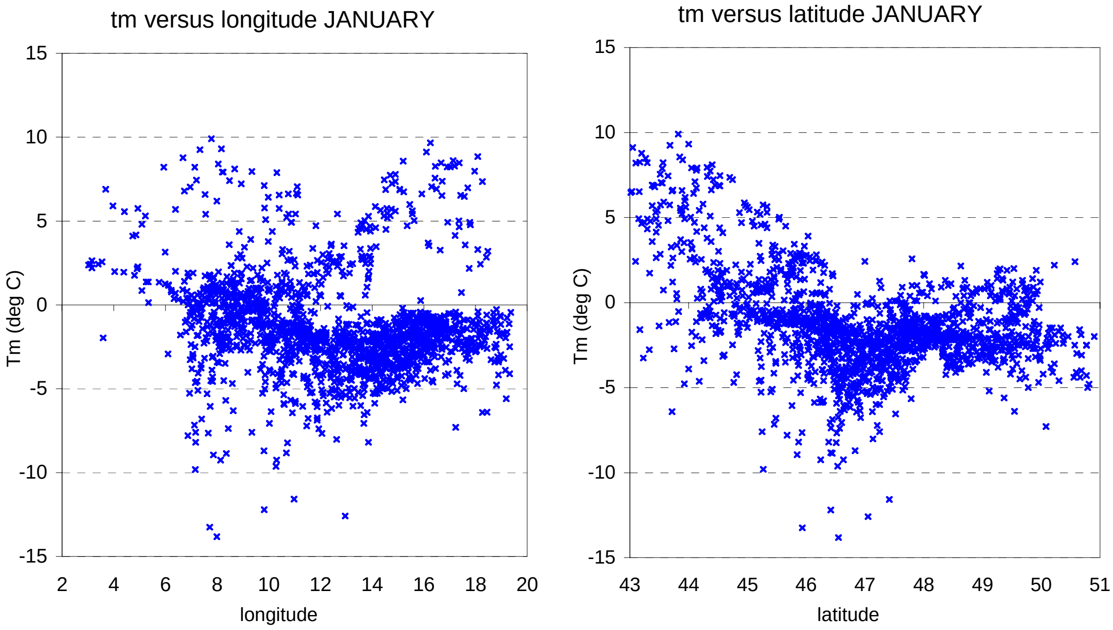


Figure15: Scatter diagram of January mean temperatures for the entire GAR versus longitude (left) and latitude (right), (sample after elimination of urban and bias-suspected stations).

Glossary of abbreviations and acronyms

CET: Central European Time
DST: Daylight Saving Time
GAR: Greater Alpine Region
MLT: Mean Local Time
tm: monthly mean temperature
tx: monthly mean maximum temperature
tn: monthly mean minimum temperature
TRM: true mean (24 hours)
AT-T24-50: Austrian hourly temperature data set of 50 stations
ZAMG: Central Institute for Meteorology and Geodynamics
HZB: Central Hydrographical Service of Austria
DTC: Daily temperature course
DTR: daily temperature range
CLINO: 30 years Climate normal period, WMO recommended
HISTALP: Historical instrumental climatological time series of the Greater Alpine Region