

# Comparison of two interpolation methods for modelling crop yields in ungauged locations

M. Dubrovsky (1), M. Trnka (2), F. Rouget (3), P. Hlavinka (2)

(1) Institute of Atmospheric Physics ASCR, Prague, Czechia ([dub@ufa.cas.cz](mailto:dub@ufa.cas.cz)),

(2) Institute of Agrosystems and Bioclimatology MUAFF, Brno, Czechia

(3) INRA CSE, Avignon, France

## *Acknowledgements:*

- Grant Agency of the Czech Republic (project 205/05/2265; [www.ufa.cas.cz/dub/calimaro/calimaro.htm](http://www.ufa.cas.cz/dub/calimaro/calimaro.htm))

- National Agency for the Agricultural Research (project QG60051)

# Abstract

The crop models linked with a weather generator are often used in assessing sensitivity of crop yields to variability and changes in climatic characteristics or in probabilistic crop yields forecasting. In these situations, weather series representing present climate, future climate or weather forecast may be produced by the weather generator, which is typically calibrated using the site-specific weather observations. The problem arises when we want to simulate crop yields for a site, where no observational data are available. Two approaches are tested in this contribution to model yields in such site: 1) Model yield is interpolated from the surrounding stations where the observed weather series are available to run the crop model. 2) The crop model is run with the site specific input parameters (including soil type) and with weather series produced by weather generator whose parameters were interpolated from the surrounding stations. In both cases, we use the nearest neighbours interpolator, in which the interpolated value (being either crop yield or weather generator parameter) is defined as a weighted average from the surrounding stations' values corrected for the mean spatial (longitudinal, latitudinal and altitudinal) linear trends. In case of the first method, the discontinuity of the spatial distribution of soil properties (which significantly affect the crop yields) is accounted for: While interpolating model yield to any target site, the model yields in the surrounding stations as well as the regression coefficients for spatial trends correction relate to the same soil type as the one found in the target site.

The two methods are tested to interpolate winter wheat yields in Czechia. We use STICS crop growth model [1] fed by daily weather series produced by Met&Roll stochastic weather generator [2, 3, 4, 5, 6, 7], which is calibrated with daily weather data from 125 stations. The soil conditions are represented by 21-soil-types map at 1x1 km resolution. Accuracy of the interpolation methods is assessed using the cross-validation technique. The set of 125 stations is split into two sets: the “learning” set includes 100 stations, from which we interpolate and which are used to derive spatial trends of variables being interpolated, and remaining 25 stations serve as “verification” stations, to which we interpolate and which are used to determine the measures of interpolation accuracy.

## Crop growth model (STICS) [1]

Crop growth models have become indispensable tools of agrometeorological and plant production research. Out of the wide range of available tools, STICS belongs to the most used and known. Although the basic philosophy of all models is the same, they differ in number of key modules and design. It is also the case of STICS model that has been built with an intention of flexible architecture allowing easy incorporation of new crop species and model functionalities. Other specific feature of the model is its consistency and transparency as the required inputs are almost exclusively in form of directly measurable parameters without use of semi-empirical coefficients.

## Weather generator (Met&Roll + GeNNeR)

The STICS model requires 6-variate daily weather series: *SRAD*, *TMAX*, *TMIN*, *PREC*, *VAPO*, *WIND*. The first four variables are produced by Met&Roll weather generator [2-7], the latter two variables (*VAPO* and *WIND*) are added using the GeNNeR generator. Met&Roll is the four-variate daily weather generator based on Markov chain model (precipitation occurrence), Gamma distribution (precipitation amount), and autoregressive model (*SRAD*, *TMAX*, *TMIN*). GeNNeR is based on the resampling procedure. Alternatively, water vapour pressure and wind speed are determined as climatological 1961-90 means for a given month (much quicker compared to using GeNNeR).

## Methods of estimating crop yields at stations with no meteo data:

**Method A:** Crop yields for the verification stations are interpolated from the surrounding learning stations. The values from the neighbourhood of user-defined size (we test 150/125/100/75/50/33/25 km surroundings) are corrected for the x-y-z trends and averaged using the bell-shaped weight function (see below)

### **notes:**

- if we want to produce the yield map for the whole territory, we must perform up to  $N_{wea} \times N_{soil}$  crop model simulations ( $N_{wea}$  = number of learning stations,  $N_{soil}$  = number of soil types in the territory).
- weather generator is not required

**Method B:** Crop yields are modelled using the site-specific input parameters (e.g. soil) and synthetic weather data produced by the interpolated weather generator (WG is interpolated from the learning stations)).

**note:** If we want to produce the yield map for the whole territory, we must perform a large number of simulations, depending on the target resolution!!!

**Interpolation:** in both approaches (interpolation of the yields in Method A; interpolation of WG in Method B), the interpolation is based on Nearest Neighbours interpolator:

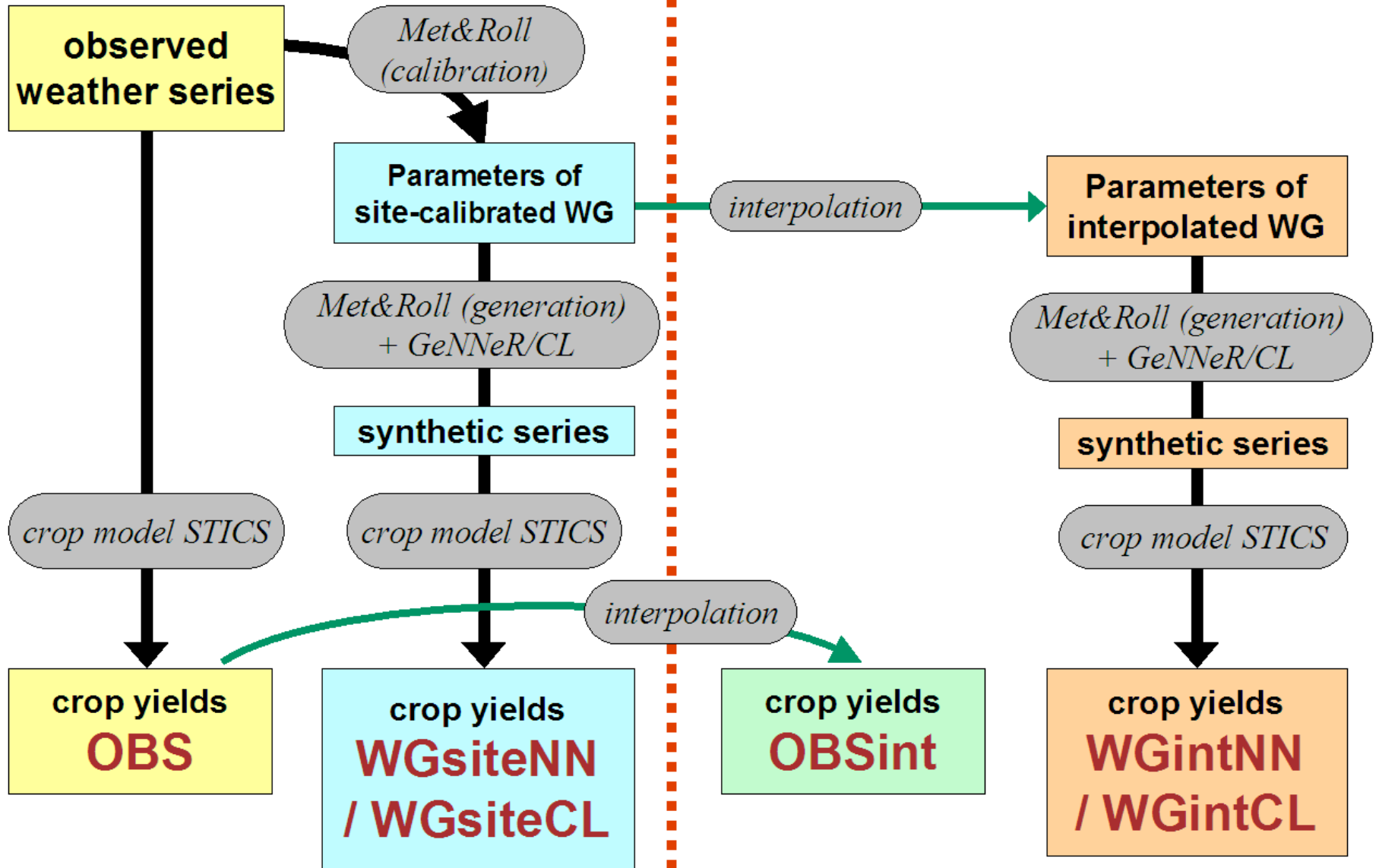
$$X^j = \sum_{i=1, \dots, N} w(D_{ij}) * X_i$$

where  $X_i$  are values in the learning stations,  $D_{ij}$  is the distance between j-th and i-th station, and  $w$  is a bell-shaped weight:

$$w(D_{ij}) = [1 - (D_{ij}/D_{max})^3]^3 \text{ for } D_{ij} \leq D_{max}, w(D_{ij}) = 0 \text{ for } D_{ij} > D_{max}$$

*set of learning stations*

*set of verification stations*



**OBS** = model yields simulated with observed weather data

**WGsiteNN** = model yields simulated with synthetic weather data produced by site-calibrated weather generator; VAPO and WIND are added by GeNNeR generator

**WGsiteCL** = model yields simulated with synthetic weather data produced by site-calibrated weather generator; VAPO and WIND are climatological means for a given month

**WGintNN** = model yields simulated with synthetic weather data produced by interpolated weather generator; VAPO and WIND are added by GeNNeR generator

**WGintCL** = model yields simulated with synthetic weather data produced by interpolated weather generator; VAPO and WIND are climatological means for a given month

**OBSint** = model yields are interpolated from the surrounding stations

**OBSint150, OBSint125, OBSint100, OBSint075, OBSint050, OBSint033, OBSint025**  
= the same as previous but using different size of the surroundings ( $D_{\max}$ ) in km

The above yield estimates were made for **three significant soil types** found in Czechia:

- **Chernosem**: the most fertile soil type in CZ
- **Cambisoil**: one of the most frequent soil type in CZ
- **Fluvisoil**: heavy soils found in the alluvial plains along the rivers

## Results - comparisons

The crop yields are compared graphically (**Figs 2a,b,c, 3,, 4**) and in **Tab.1**.

**(WGsiteNN vs OBS) and (WGsiteCL vs OBS)**: the differences are due to imperfections of weather generator and stochasticity involved in the generator

**(WGsiteNN vs WGsiteCL), (WGintNN vs WGintCL)**: the differences are due to different approach in generating humidity and wind. The differences are found negligible (indicated by high correlations).

**(WGintNN vs WGsiteNN) and (WGintCL vs WGsiteCL)**: the differences are due to errors in interpolating the generator: the results show quite high fit (indicated by high correlations)

**... and the most important:**

**OBSint vs OBS**: the differences show the performance of method A.

**WGintXX vs OBS**: the differences show the performance of the method B.

## Conclusions

- 1) In generating the synthetic series, the Met&Roll+CL method, which is much faster than the Met&Roll+GeNNeR method provides nearly the same results.
- 2) In interpolating the yields, 125 km surroundings provides best results.

**3) Of the two methods used to estimate crop yields for the stations with no weather data, the Method B, which involves interpolated weather generator shows much higher performance than the Method A.**

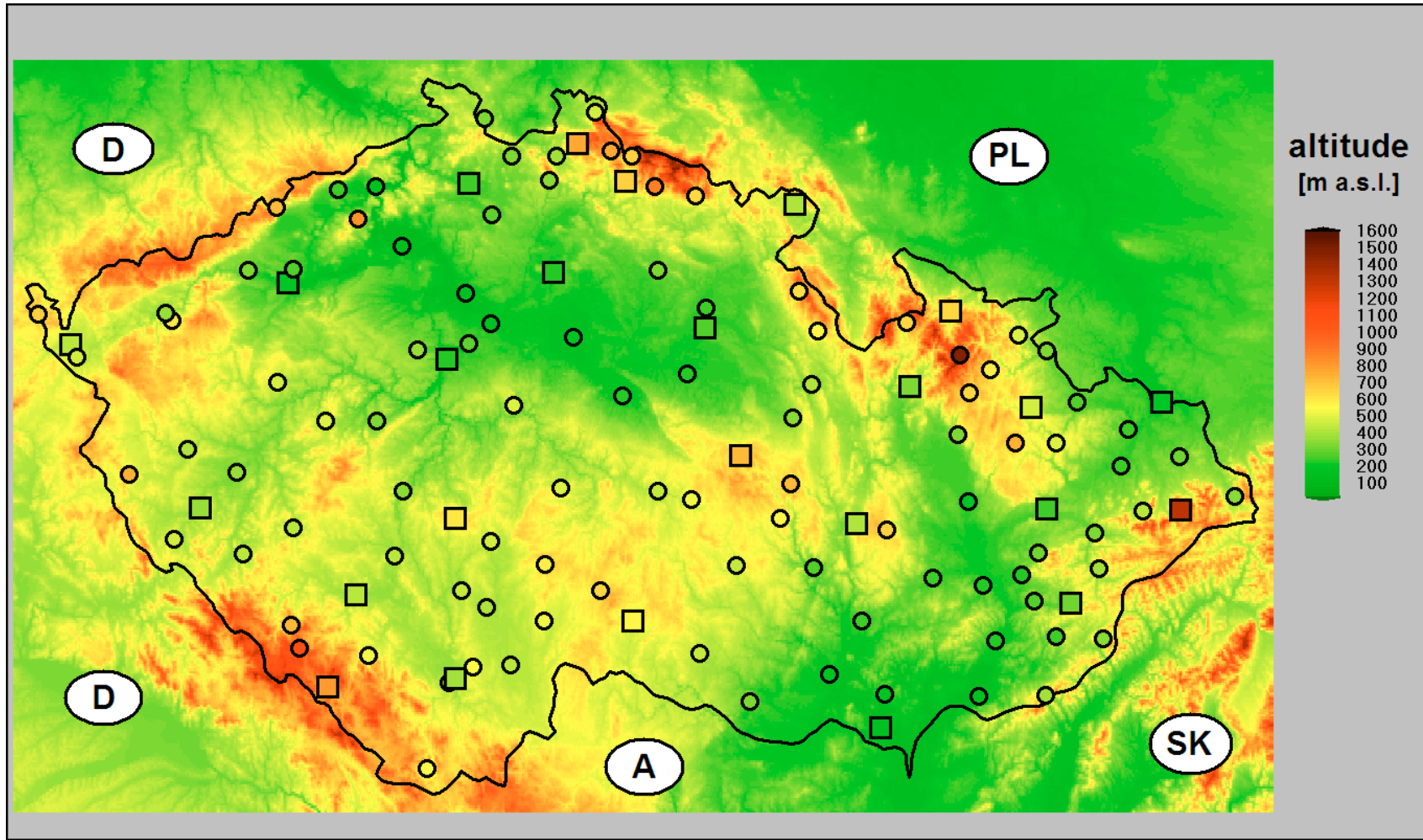
## References:

- [1] Brisson N. et al., 2003, Europ. J. Agronomy, 18, 309-332
- [2] Dubrovsky M. et al, 2000, Clim. Change 46, 447-472
- [3] Zalud Z. and Dubrovsky M., 2002, Theor. Appl. Climatol., 72, 85-102;
- [4] Dubrovsky M. et al., 2004, Clim. Change 63, 145-179;
- [5] Trnka M. et al., 2004, Theor. Appl. Climatol. 77, 229-249
- [6] Trnka M. et al., 2004, Clim. Change 64, 227-255
- [7] Kysely J. and Dubrovsky M., 2005, Int.J.Climatol. 25, 251-269.

---

the poster is available at [www.ufa.cas.cz/dub/calimaro/calimaro.htm](http://www.ufa.cas.cz/dub/calimaro/calimaro.htm)

# Fig.1: Region (= Czechia) and Data



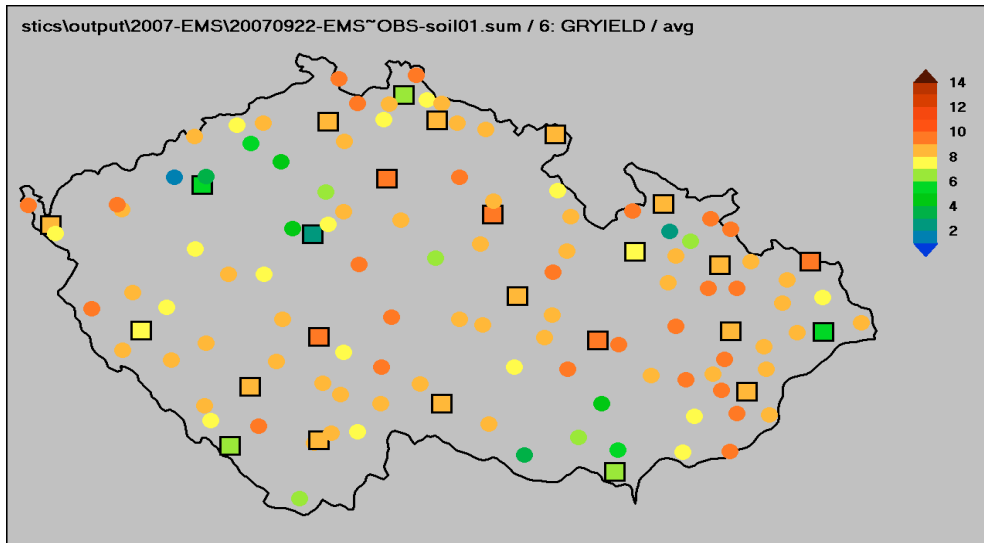
**30y weather series (1961-1990) are available from 125 Czech stations:**

- **circles:** “learning” set, **squares:** “validation” set

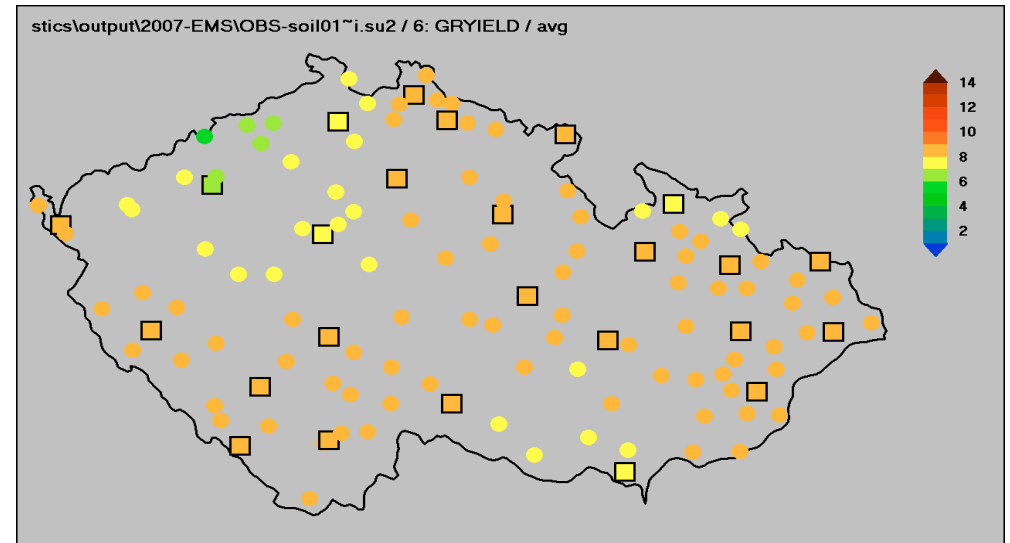
**altitude** of Czechia varies from 115 to 1602 m a.s.l. (topography derived from the global digital elevation model GTOPO30 /horizontal grid spacing = 30 arc seconds/; <http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html>)

# Fig. 2a: Mean model wheat yields [soil = Chernozem (CZ\_01)]

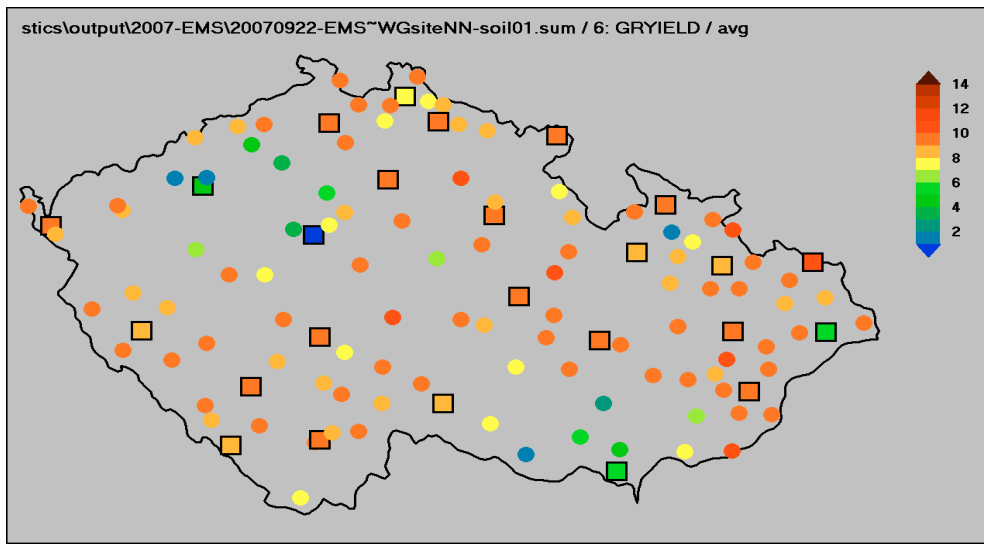
[the learning stations are marked by circles, verification stations by squares]



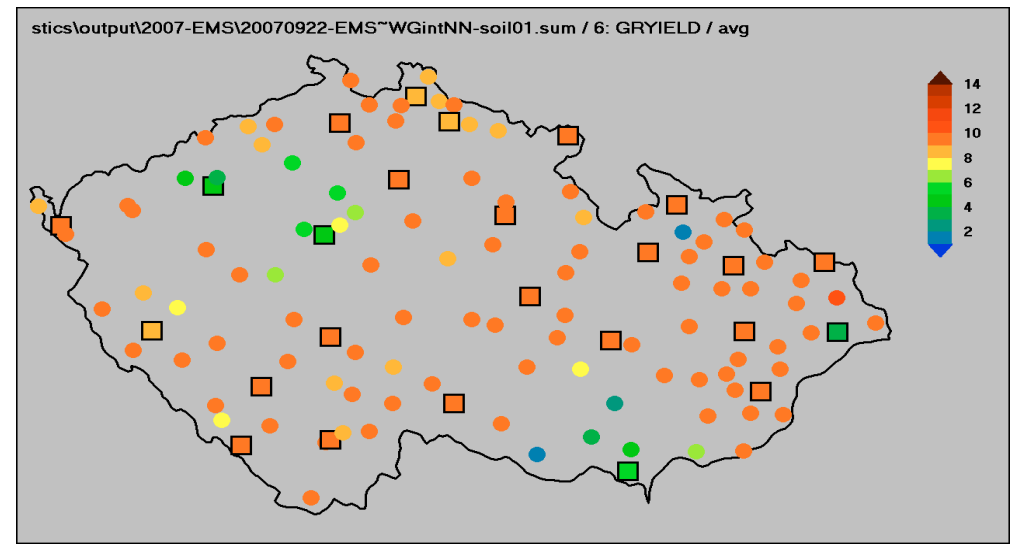
a) weather = observed weather series



b) yields interpolated from the set of learning stations



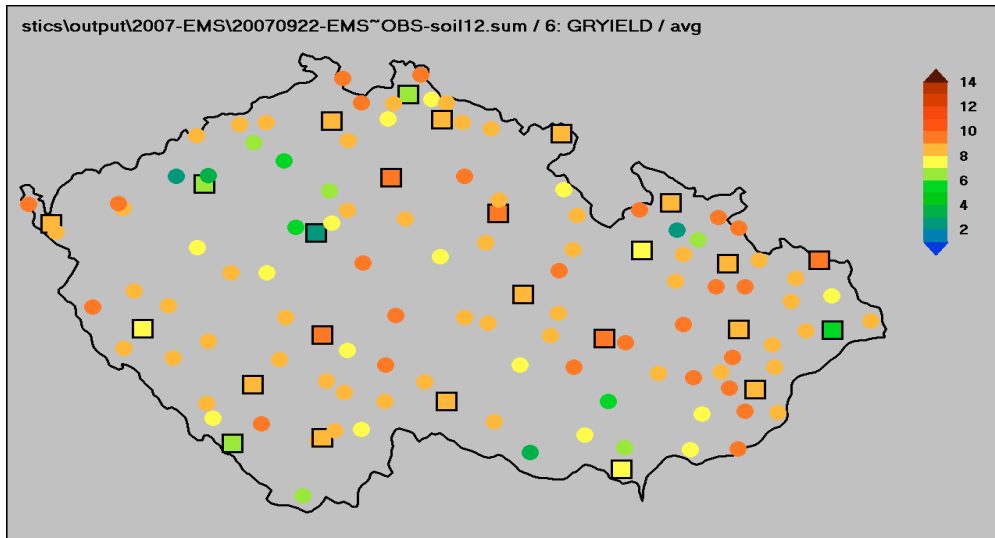
c) synthetic weather produced by site-specific WG  
(site-calibrated Met&Roll + VAPO and WIND added by GeNNer)



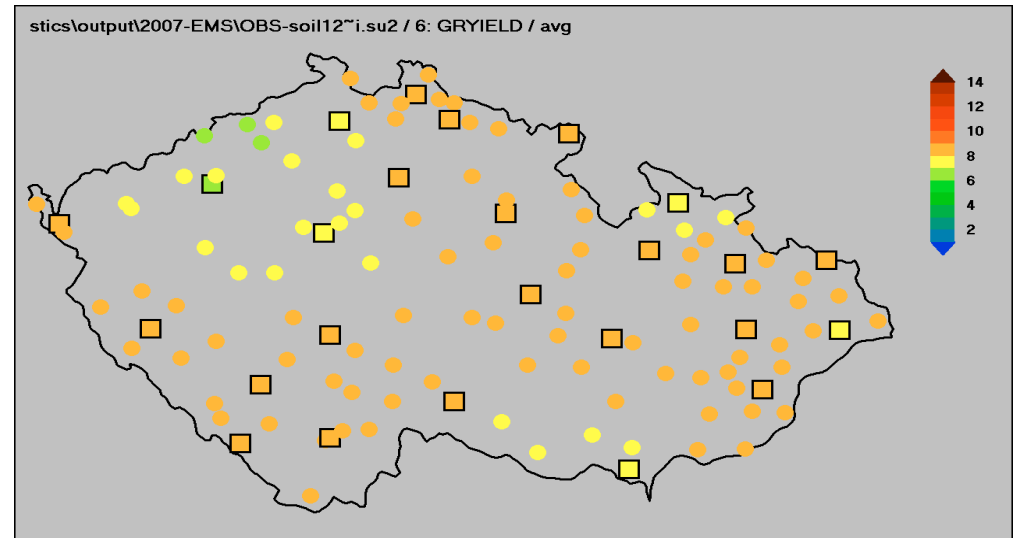
d) synthetic weather produce by interpolated WG  
(interpolated Met&Roll + VAPO, WIND added by GeNNer)

# Fig.2b: Mean model wheat yields [soil = Cambisoi (CZ\_12)]

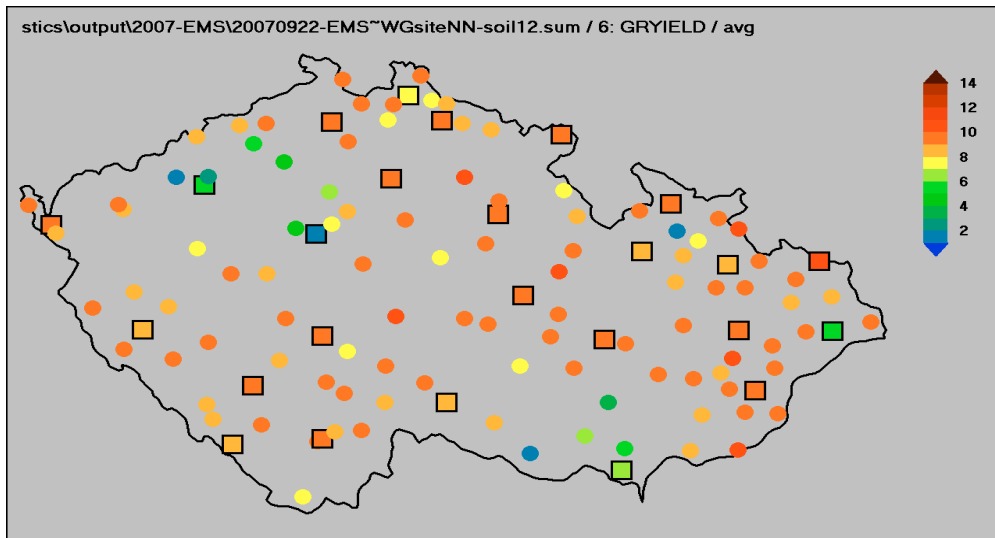
[the learning stations are marked by circles, the verification stations by squares]



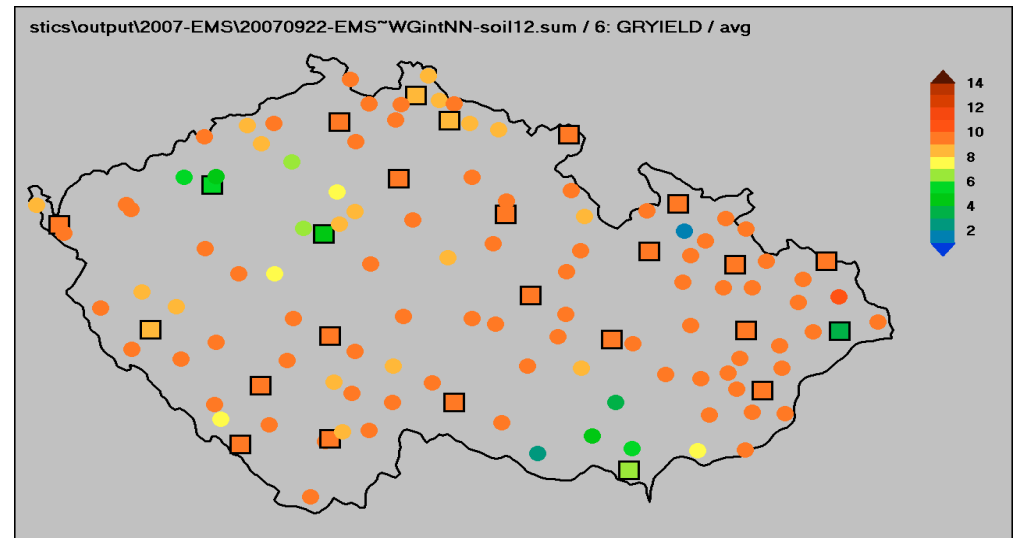
a) weather = observed weather series



b) yields interpolated from the set of learning stations



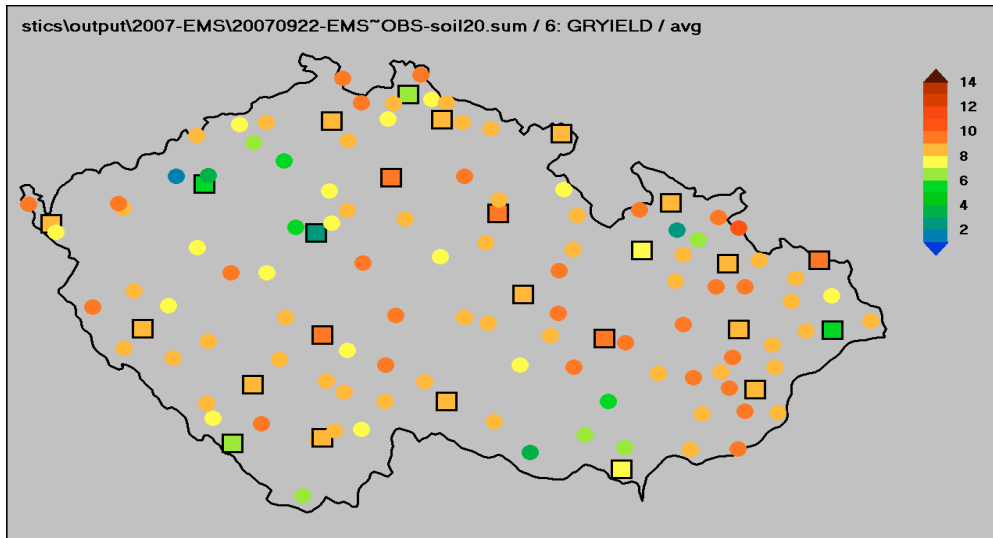
c) synthetic weather produced by site-specific WG  
(site-calibrated Met&Roll + VAPO and WIND added by GeNNer)



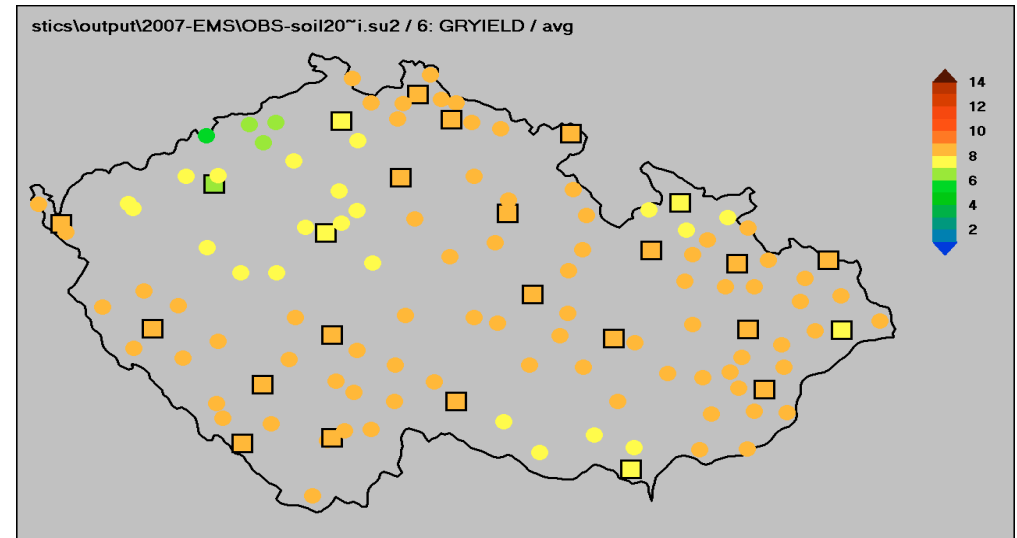
d) synthetic weather produce by interpolated WG  
(interpolated Met&Roll + VAPO, WIND added by GeNNer)

# Fig.2c: Mean model wheat yields [soil = Fluvisoil (CZ\_20)]

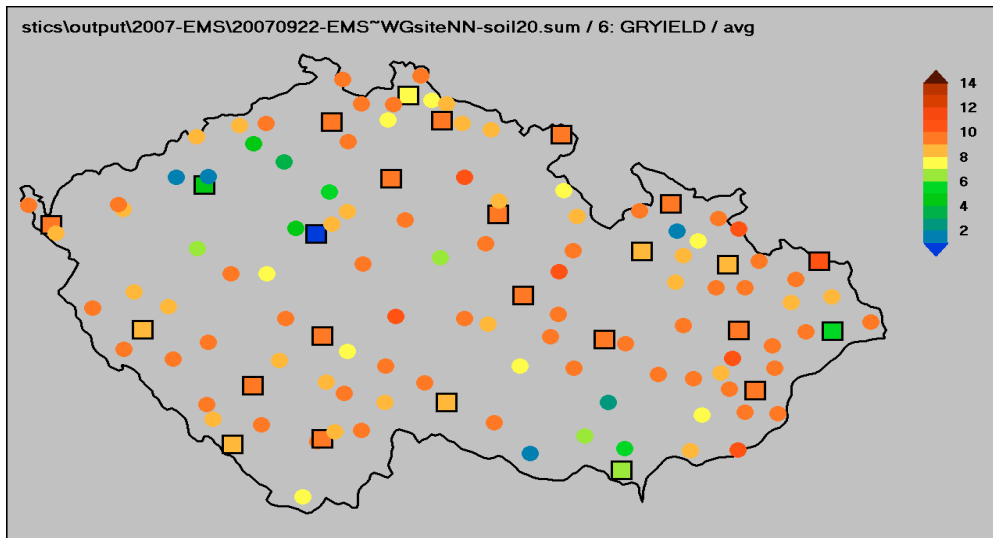
[the learning stations are marked by circles, the verification stations by squares]



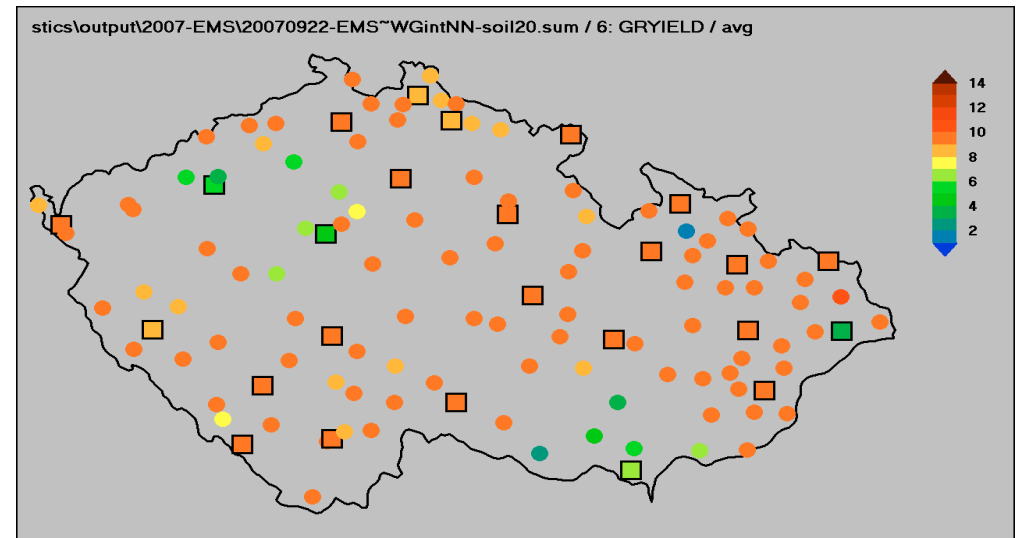
a) weather = observed weather series



b) yields interpolated from the set of learning stations



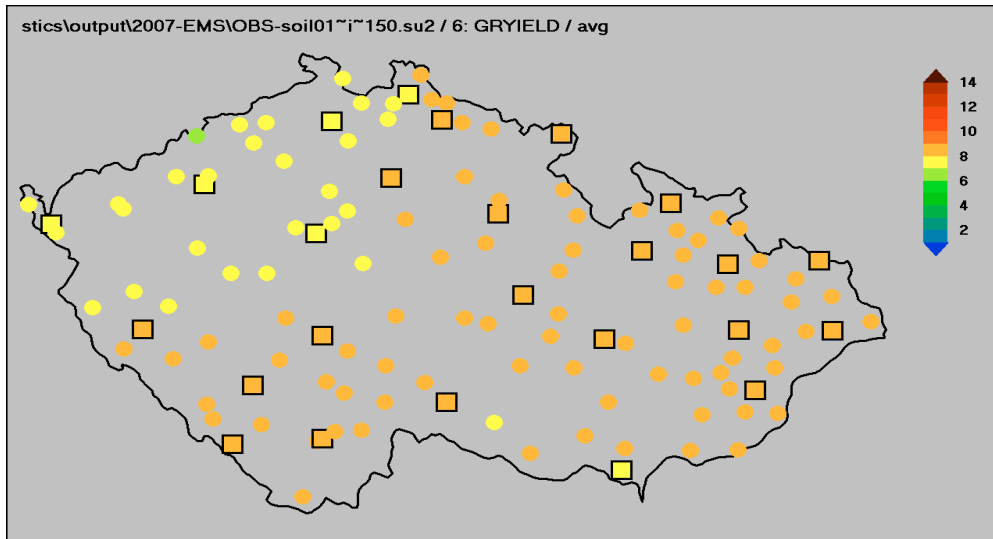
c) synthetic weather produced by site-specific WG  
(site-calibrated Met&Roll + VAPO and WIND added by GeNNer)



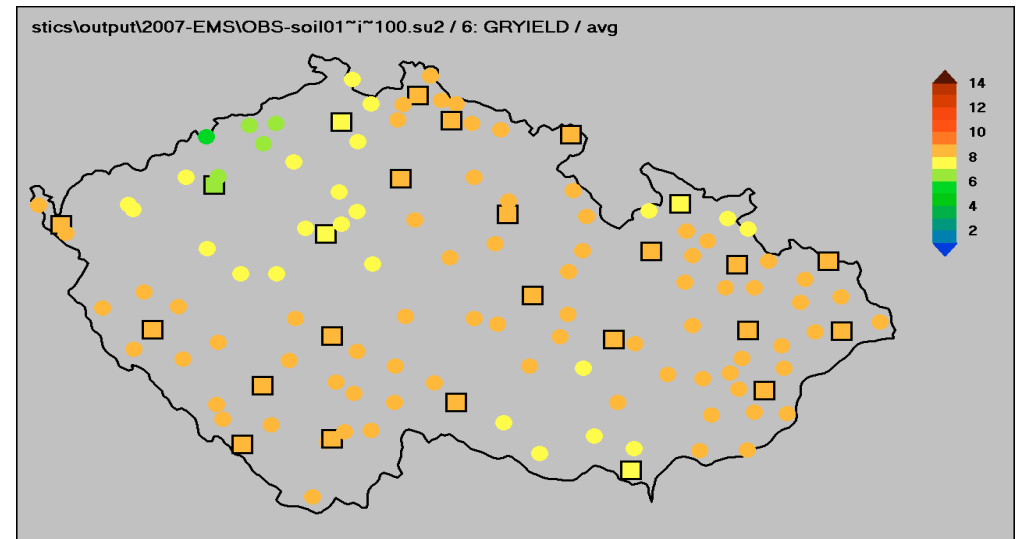
d) synthetic weather produce by interpolated WG  
(interpolated Met&Roll + VAPO, WIND added by GeNNer)

# Fig.3: Method A: effect of the neighbourhood size

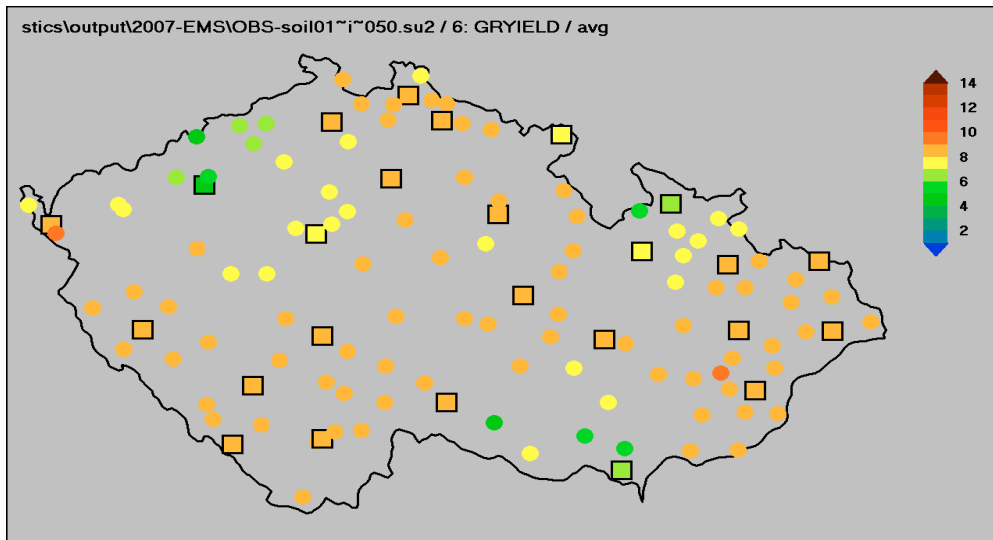
[soil = Chernozem(CZ\_01); circles: learning stations; squares = verification stations]



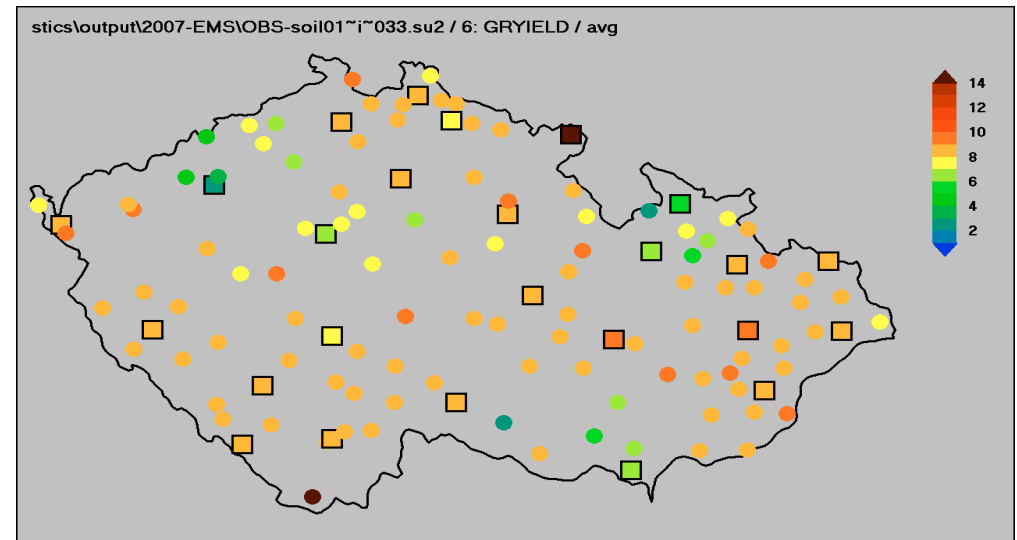
a) interpolated yields;  $D_{\max} = 150$  km



b) interpolated yields;  $D_{\max} = 100$  km

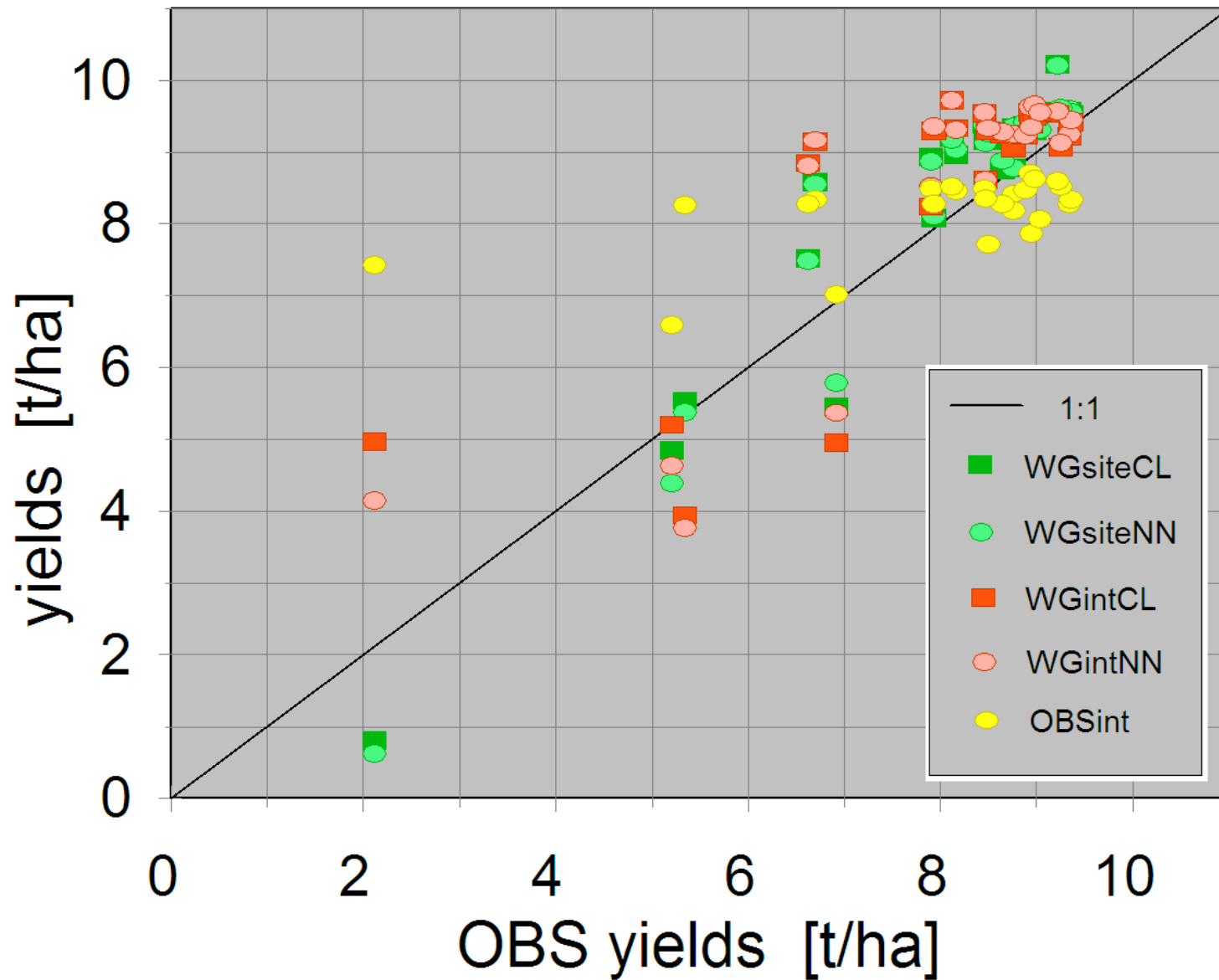


c) interpolated yields;  $DIST_{\max} = 50$  km



d) interpolated yields;  $DIST_{\max} = 33$  km

**Fig.4: Validation of the interpolation methods**  
[wheat yields in the verification stations; soil = Chernozem]



# Tab. 1: validation of the interpolation methods

(the correlations are calculated only from the verification stations)

	Chernozem CZ_01	Cambisoil (CZ_12)	Fluvisoil (CZ_20)
cor(WGsiteNN, OBS)	0.961	0.951	0.951
cor(WGsiteCL, OBS)	0.955	0.952	0.948
cor(WGintNN, OBS)	0.851	0.827	0.825
cor(WGintCL, OBS)	0.810	0.791	0.791
cor(OBSint, OBS)	0.581	0.590	0.574
cor(WGsiteCL, WGsiteNN)	0.998	0.998	0.999
cor(WGintCL, WGintNN)	0.993	0.991	0.995
cor(WGintNN, WGsiteNN)	0.905	0.885	0.892
cor(WGintCL, WGsiteCL)	0.872	0.842	0.863
cor(OBSint150, OBS)	0.501	0.516	0.531
cor(OBSint125, OBS)	0.581	0.590	0.603
cor(OBSint100, OBS)	0.584	0.570	0.574
cor(OBSint075, OBS)	0.550	0.529	0.518
cor(OBSint050, OBS)	0.536	0.488	0.481
cor(OBSint033, OBS)	0.540	0.486	0.481
cor(OBSint025, OBS)	0.500	0.440	0.446

---

the poster is available at  
[www.ufa.cas.cz/dub/calimaro/calimaro.htm](http://www.ufa.cas.cz/dub/calimaro/calimaro.htm)

---