Using Kriging Combined with Categorical Information of Soil Maps for Interpolating Soil Properties

Dar-Yuan Lee*[1], Kai-Wei Juang[2], and Ten-Lin Liu[3]

[1] Department of Agricultural Chemistry, National Taiwan University

[2] Department of Agronomy, National Chiayi University

[3] Tea Research and Extension Station, Taiwan

Introduction

- Kriging interpolation is frequently used for mapping soil properties in the analysis and interpretation of spatial variation of soil.
- Mapping quality could affect the performance of site-specific management.
- Soil map–delineation in existing soil maps showing abrupt changes at the boundaries between different soil types can provide valuable categorical information for interpreting variation in soil properties.

Introduction (continued)

- In this study, map units were used to group sampled observations, and the variation in soil properties was separated into two parts: (i) between soil types (i.e., soil type effect) and (ii) within each soil type (i.e., residual).
- A kriging model combined with the information of soil map–delineation, taking into account the variation components of soil type effect and residual, was proposed.
- A comparison of performance of kriging combined with soil map–delineation (KSMD) and ordinary kriging (OK) was performed to assess the feasibility of KSMD for improving the interpolation of soil properties.

Methodology

- The notation of (A_j) was used to quantify the information of soil map–delineation as nominal data. The observations are denoted with the same soil type (j). Thus, each observation $z_j(x_i)$ on one specific soil property can be expressed as

$$\boxed{z(x_{ij}) = s(A_j) + r(x_{ij}),}$$

where **x_{ij} is the sampling location** of $z(x_{ij})$, which is on the polygon A_j. **$s(A_j)$ is the mean value of $z(x_{ij})$ over the polygon A_j, and $r(x_{ij})$ is the residual**.

Methodology (continued)

- As $s(A_j)$ and $r(x_{ij})$ are independent mutually and the variation of $r(x_{ij})$ is assumed to be homogeneous over all polygons. The variance σ^2_z of $z(x_{ij})$ can be shown as:

$$\sigma^2_z = \sigma^2_s + \sigma^2_r$$

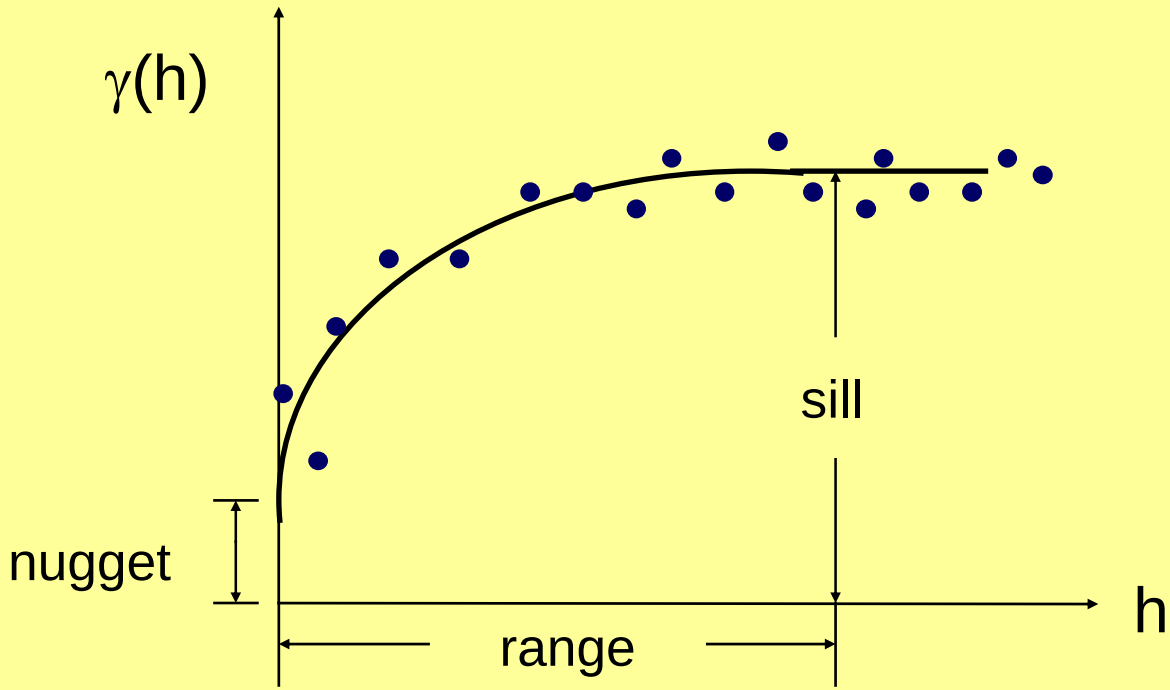
where the variance σ^2_s **indicates the spatial variation between polygons**. The variance σ^2_r **shows the overall variation within polygons**, which is defined homogeneously as **the intrinsic stationarity is assumed for kriging estimation**.

Methodology (continued)

- If the variance σ^2_z of $z(x_{ij})$ is not as stationary as σ^2_r showing the variation of $r(x_{ij})$ within polygons, **kriging estimation will be more suitable on the variable domain of $r(x_{ij})$ than that of $z(x_{ij})$.**

- A new approach, **Kriging combined with soil map-delineation** (**KSMD),** for spatial interpolation of soil properties was proposed.

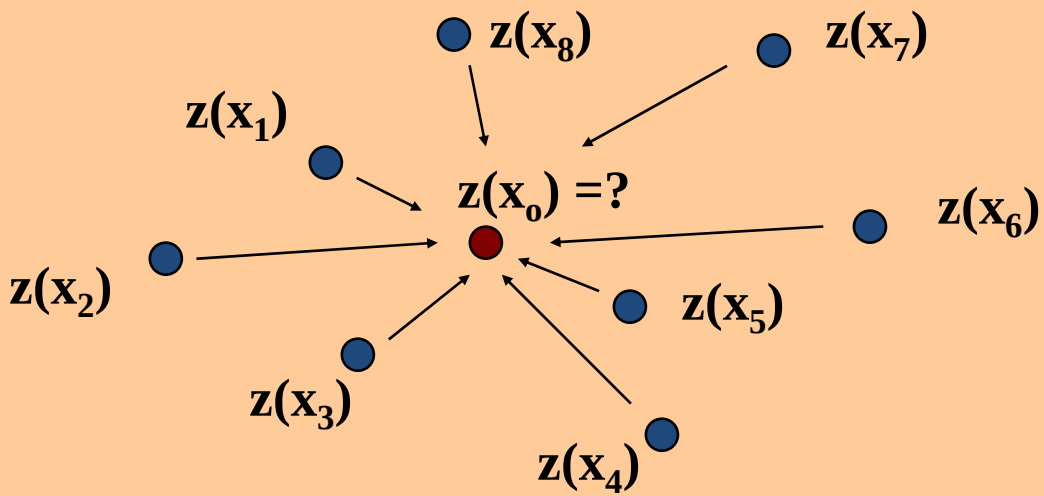
Commonly used interpolation: Ordinary kriging



Semivariance: $\gamma(h) = ½ \, Var[z(x+h) - z(x)]$

- $$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i+h) - z(x_i)]^2$$

Commonly used interpolation: Ordinary kriging (continued)



$$z^*(x_o) = z(x_1)\lambda_1 + z(x_2)\lambda_2 + z(x_3)\lambda_3 + z(x_4)\lambda_4 + z(x_5)\lambda_5 + z(x_6)\lambda_6 + z(x_7)\lambda_7 + z(x_8)\lambda_8$$

Kriging combined with soil map–delineation (KSMD) vs. Ordinary kriging (OK)

Ok estimations

KSMD estimations

The original value $Z(x)$ of soil properties separating into two parts $s(A)$ and $r(x)$

$$z(x_{kj}) = s(A_j) + r(x_{kj})$$

Theoretical assumption

The ordinary kriging of the original values

$$z^*(x_o) = \sum_{i=1}^{n} \lambda_i z(x_i)$$

OK

The mean value of each soil type $s(A_j)$

The residual $r(xj)$

The ordinary kriging of the residuals

$$r^*(x_{ol}) = \sum_{j=1}^{m} \sum_{k=1}^{n(j)} \lambda_{kj} r(x_{kj})$$

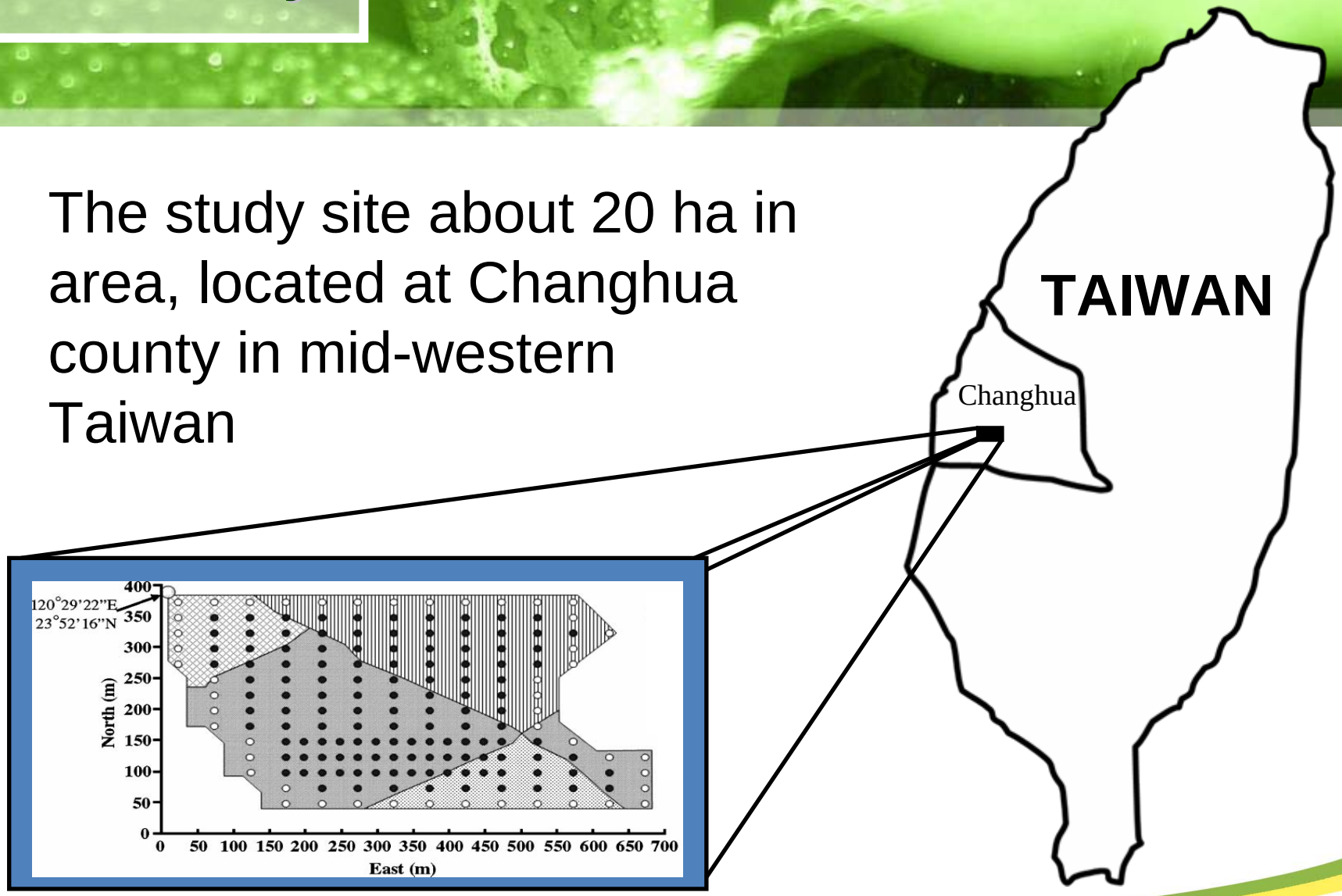
The estimator of kriging combined with soil map-delineation (KSMD)

$$z^*(x_{ol}) = s^*(A_l) + r^*(x_{ol})$$

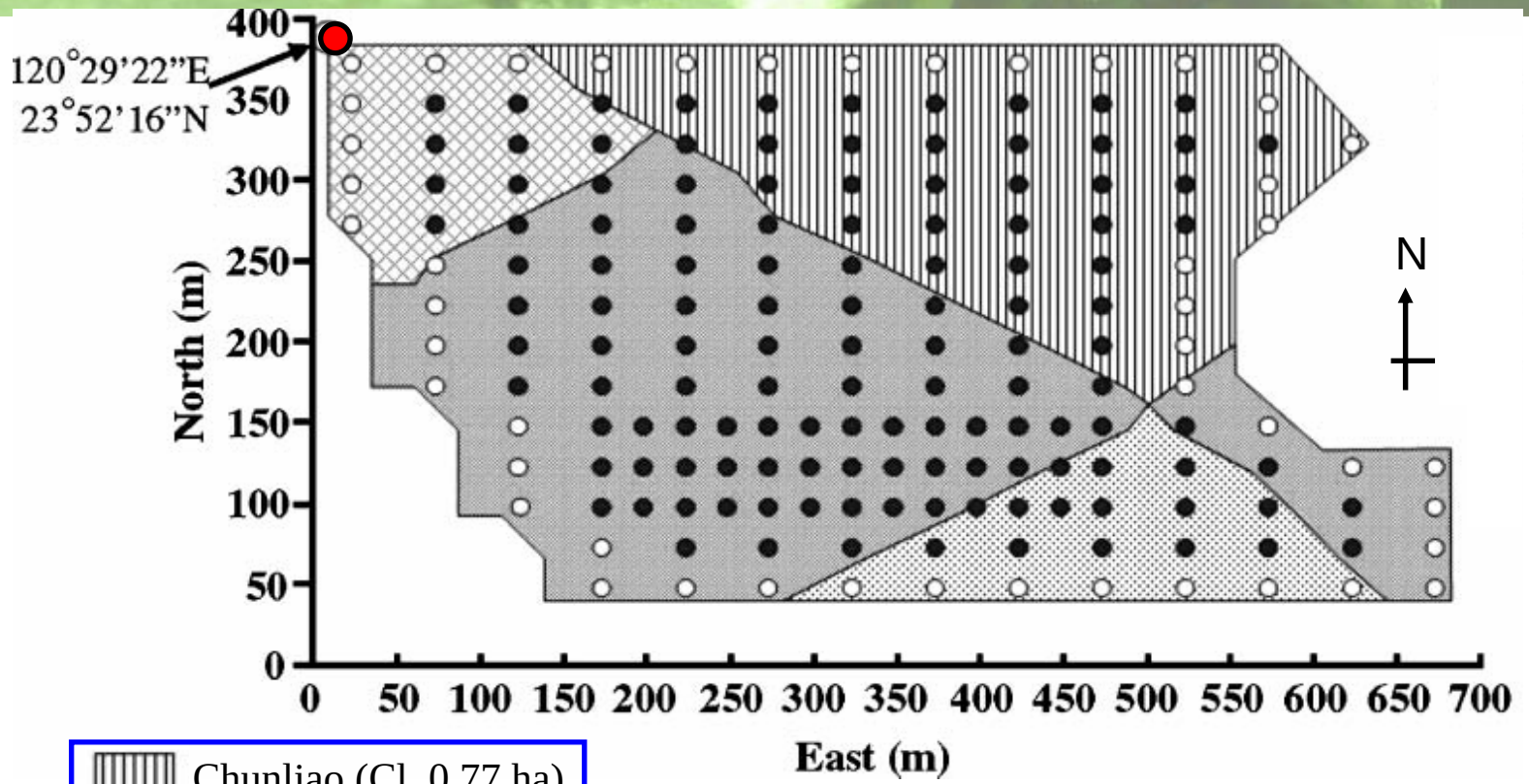
KSMD

Case study

The study site about 20 ha in area, located at Changhua county in mid-western Taiwan



Soil map classification and sampling configuration



Chunliao (Cl, 0.77 ha)

Tingfanp (Ti, 3.86 ha)

Wanhoh (Wa, 7.15 ha)

Yuliao (Yl, 2.47 ha)

Sampling locations: 175

○ 48 locations for sampling but not for validation

● 127 locations both for sampling and validation

Statistics for soil texture

Soil type	n	Sand			Silt			Clay		
		Mean	SD†	SE‡	Mean	SD	SE	Mean	SD	SE
		%								
Chunliao	49	44.7b§	9.9	0.20	35.2b	7.3	0.15	20.1bc	3.4	0.07
Tingfanpo	21	36.8a	13.0	0.62	41.7c	10.0	0.48	21.5c	4.0	0.19
Wanhoh	18	58.8c	9.1	0.51	24.3a	6.1	0.34	16.9a	3.7	0.21
Yuliao	87	32.8a	9.9	0.11	43.1c	7.5	0.09	24.1d	4.0	0.05

† Standard deviation.

‡ Standard error (SE = SD/n).

§ Mean values followed by different lowercase letters are significantly different ($P < 0.05$) between soil types.

Statistics for pH, M3-P, and M3-Ca

		pH			Mehlich-3 P			Mehlich-3 Ca		
Soil type	*n*	Mean	SD	SE	Mean	SD	SE	Mean	SD	SE
					mg kg^{-1}			mg g^{-1}		
Chunliao	49	7.8b	0.2	0.01	44.4a	23.2	0.47	2.8d	1.0	0.02
Tingfanpo	21	7.4a	0.8	0.04	52.1a	15.0	0.71	1.4b	0.6	0.03
Wanhoh	18	7.1a	0.7	0.04	46.1a	37.2	2.01	0.8a	0.5	0.03
Yuliao	87	7.8b	0.2	0.00	40.8a	17.2	0.20	1.9c	0.6	0.01

† Standard deviation.

‡ Standard error (SE = SD/*n*).

§ Mean values followed by different lowercase letters are significantly different ($P < 0.05$) between soil types.

Analysis of variance (ANOVA) for soil type effects

Soil property	Source of variance	*df*	Sum of squares	Mean square	*F* value
Sand, %	Soil type effect	3	12 156.22	4052.07	38.79**
	Residual (within soil types)	171	17 863.43	104.46	
	Total	174	30 019.66		
Silt, %	Soil type effect	3	6 232.97	2077.66	35.21**
	Residual (within soil types)	171	10 090.93	59.01	
	Total	174	16 232.90		
Clay, %	Soil type effect	3	1 043.85	347.95	24.17**
	Residual (within soil types)	171	2 461.36	14.39	
	Total	174	3 505.21		
pH (1:1 H_2O)	Soil type effect	3	10.22	3.41	21.77**
	Residual (within soil types)	171	26.76	0.16	
	Total	174	36.98		
Mehlich-3 P, mg kg^{-1}	Soil type effect	3	2 365.23	788.41	1.71
	Residual (within soil types)	171	79 058.54	462.33	
	Total	174	81 423.77		
Mehlich-3 Ca, mg kg^{-1}	Soil type effect	3	66.52	22.17	41.52**
	Residual (within soil types)	171	91.33	0.53	
	Total	174	157.85		

** Significant at the 0.01 level.

Soil type effect

The results of ANOVA showed that soil type effects were significant on soil texture, pH, and M3-Ca, but not on M3-P.

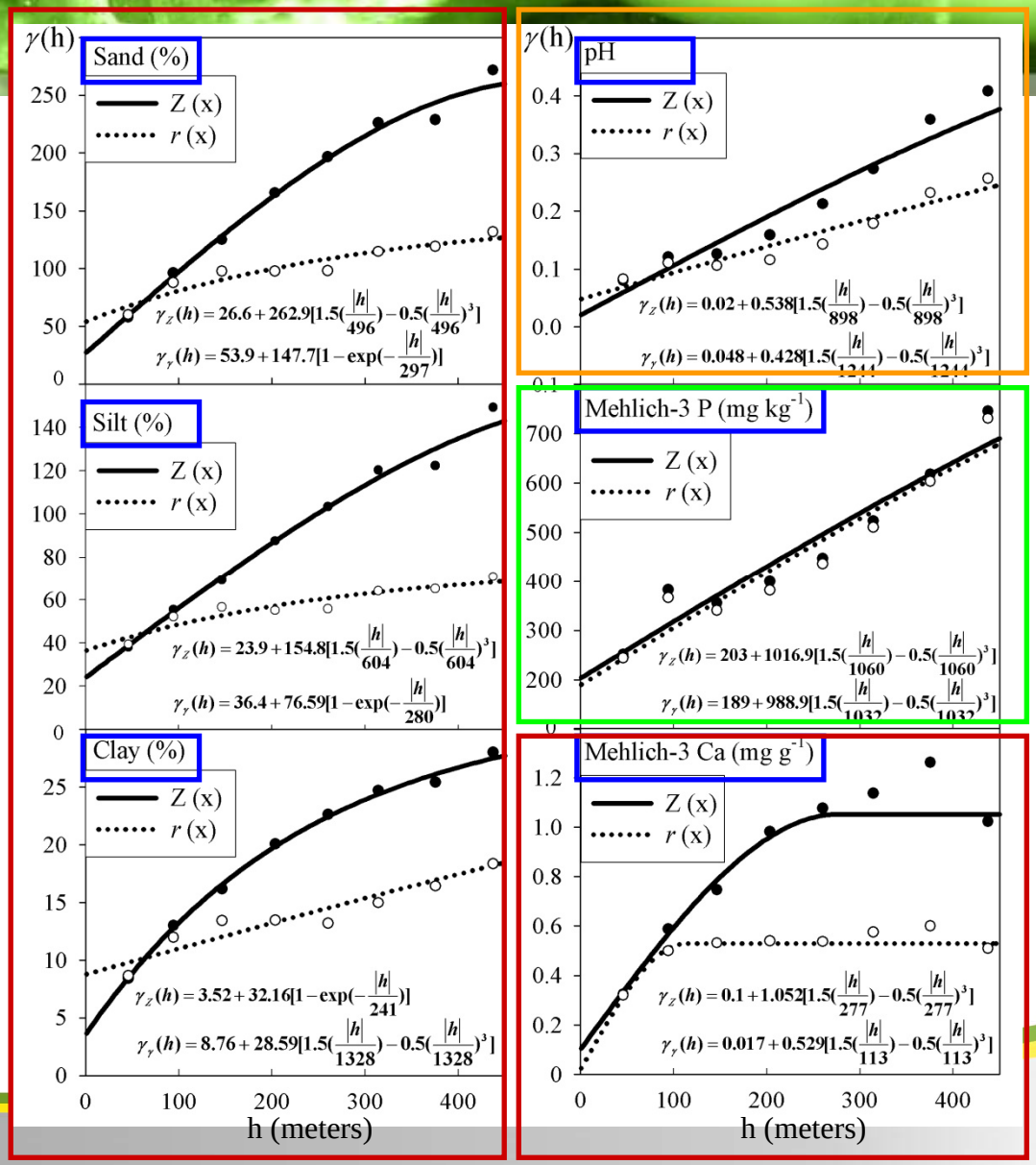
Because soil texture, pH and soil calcium are dominated by geological and pedological features, so the variation of their observed data is significantly related to soil map-delineation.

In contrast, the soil P is usually influenced by historical practices and fertilizer applications. Thus, the soil map-delineation is not significant to the data of M3-P.

Variograms for soil texture, pH, M3-P, and M3-Ca

Original value: $Z(x)$

Residual: $r(x)$



Spatial structures of residuals vs. original values for M3-P

- The variograms of residuals $r(x_{ij})$ and original values $z(x_{ij})$ for M3-P were much the same.
- This indicated that soil types were not the main sources of the spatial variation of M3-P.
- This phenomenon was consistent with the ANOVA testing and shows that the mean values $s(A_i)$ of M3-P between soil types were not different significantly.

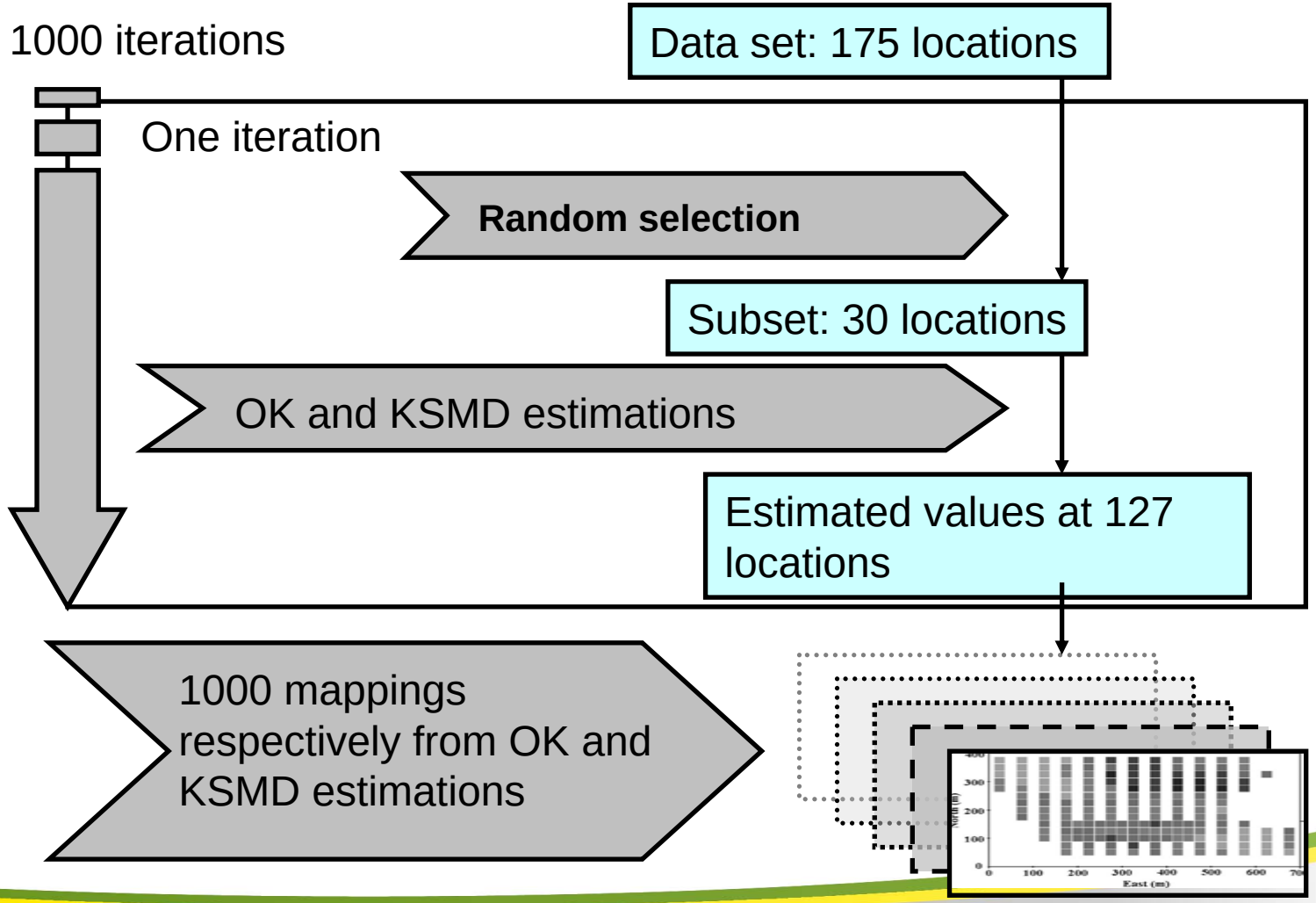
Spatial structures of residuals vs. original values for soil texture, pH, and M3-Ca

For soil texture, pH, and M3-Ca, the variograms showed that the spatial structures of the residuals $r(x_{ij})$ were obviously different from those of the original values $z(x_{ij})$.

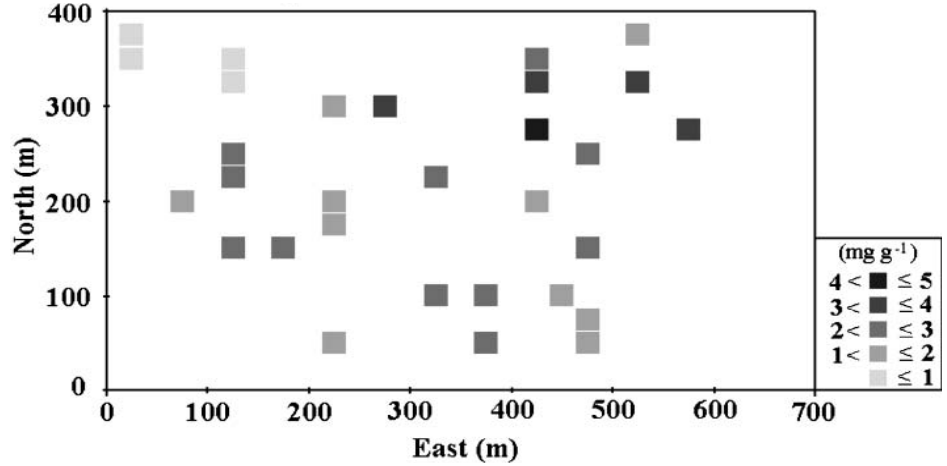
This indicates that the data of these soil properties were not stationary over the whole site. The contribution of the mean values $s(A_j)$ of the soil types (called a local trend effect) to the spatial variation of $z(x_{ij})$ was sufficiently substantial to influence the kiging estimation.

It is necessary to come up with a derived method that takes into account the local trend effect (i.e. soil type effect) for spatial interpolation of these soil properties.

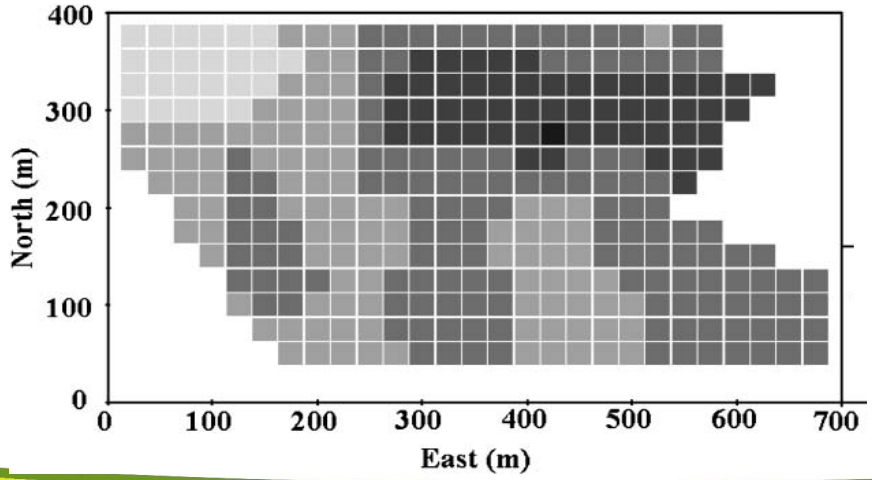
Simulation procedure for OK and KSMD estimations



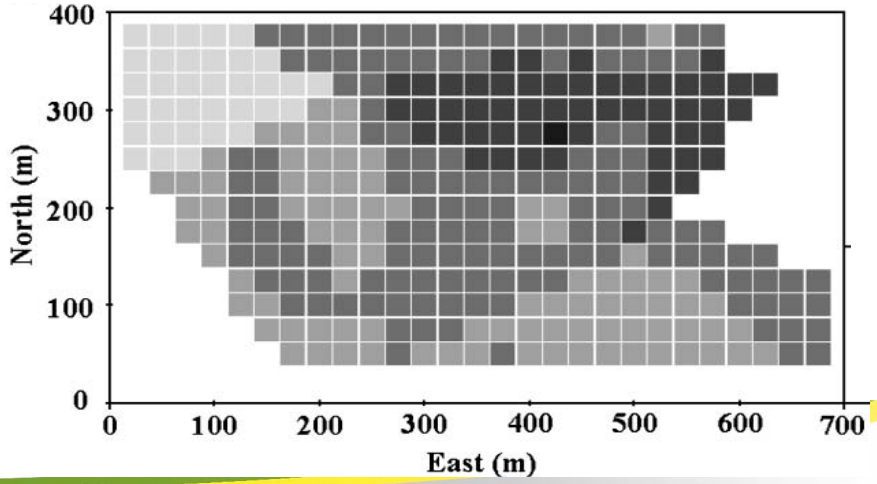
One realization for 30 sample values of M3-Ca randomly drawn from the 175 observations



OK mapping



KSMD mapping



Comparison of OK and KSMD estimations

Mean error:

$$ME(x_o) = \frac{1}{1000}\sum_{i=1}^{1000}[Z^*_i(x_o) - Z(x_o)]$$

Root mean squared error:

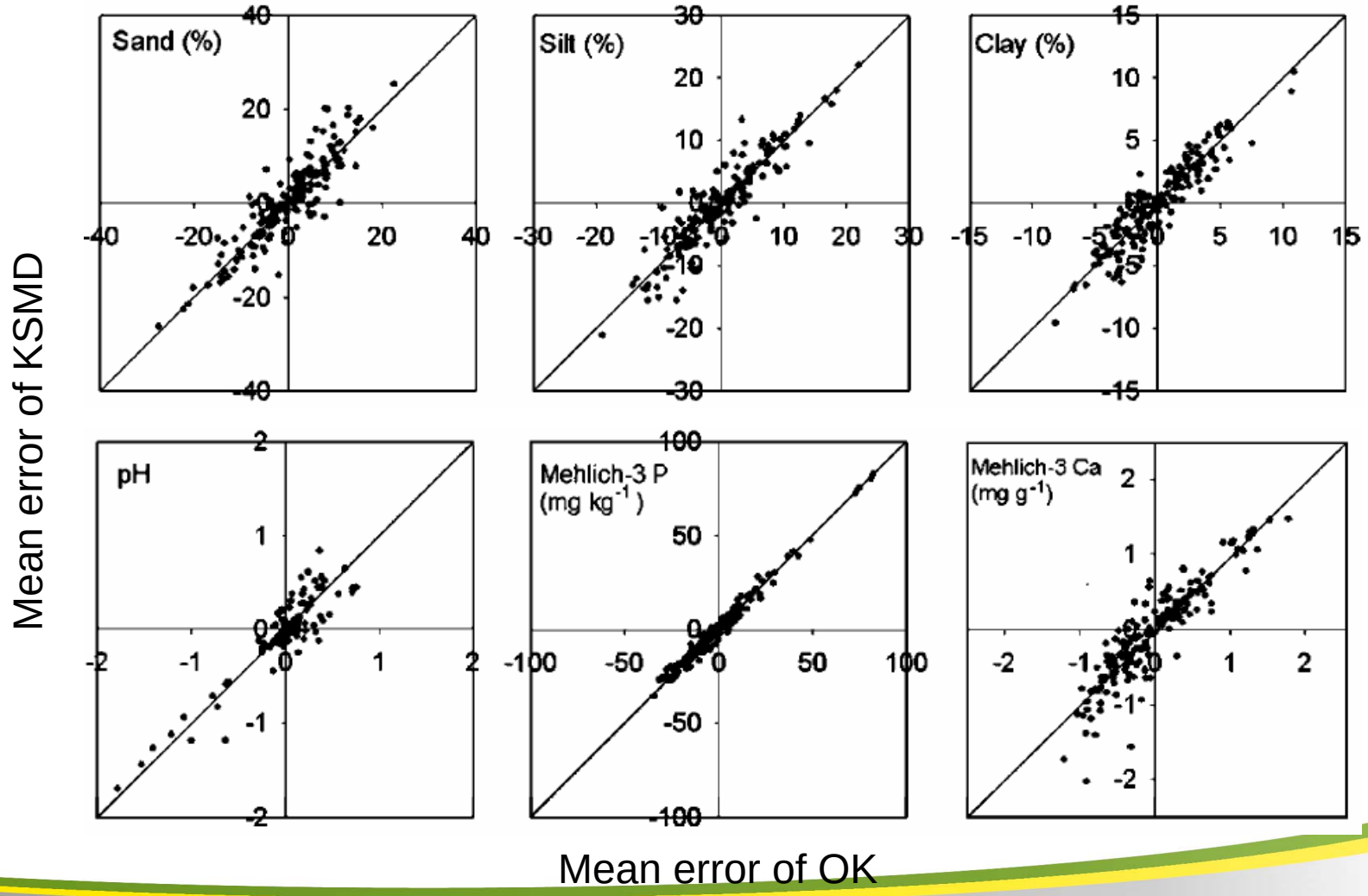
$$RMSE(x_o) = \left\{\frac{1}{1000}\sum_{i=1}^{1000}[Z^*_i(x_o) - Z(x_o)]^2\right\}^{1/2}$$

Imprecision index:

$$IP^2(x_o) = RMSE^2(x_o) - ME^2(x_o)$$

X_o is each of the 127 locations and i is denoted as the iteration number.

Accuracies of OK and KSMD estimations for soil texture, pH, M3-P, and M3-Ca

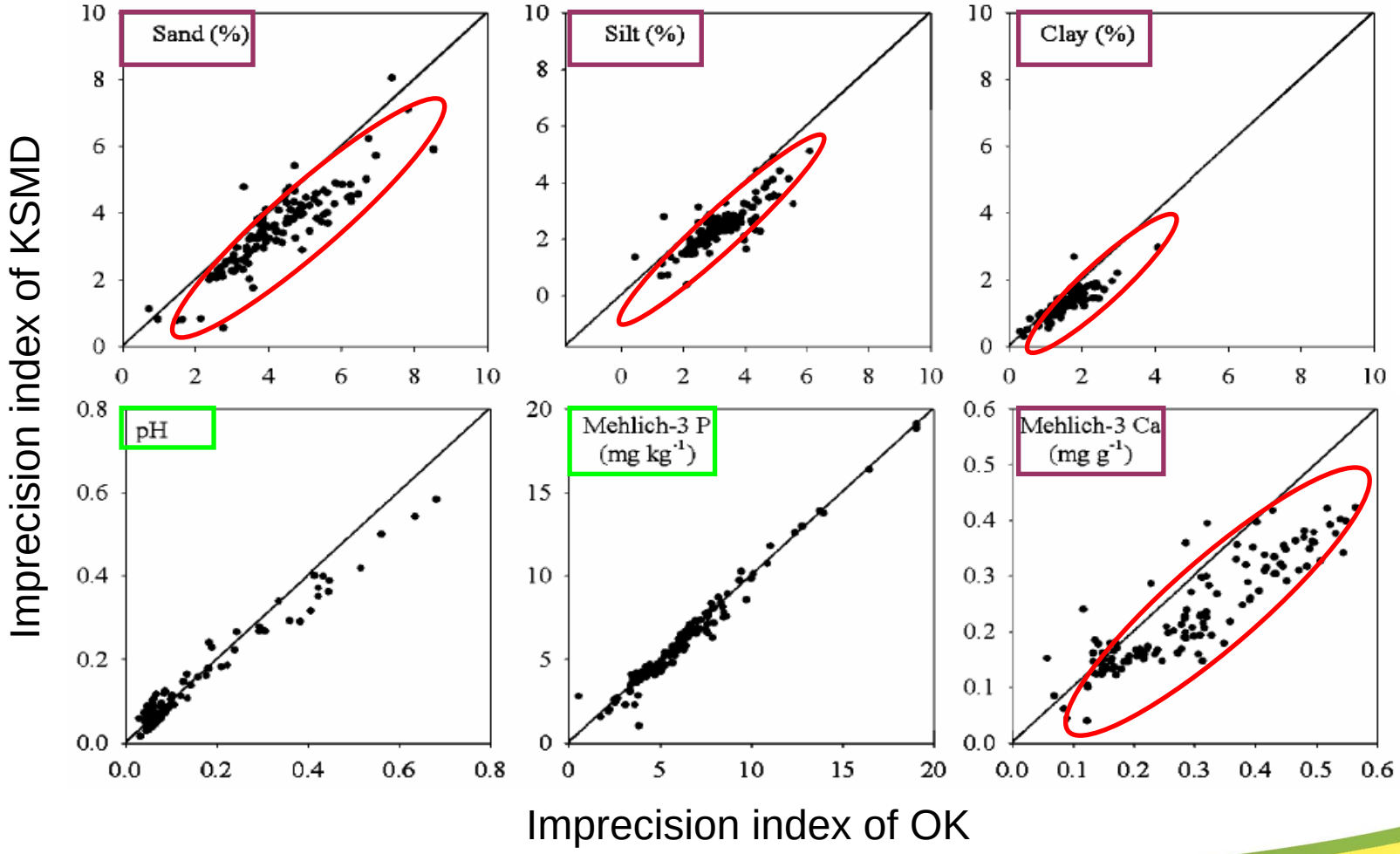


Comparison of accuracies of OK and KSMD estimations

In comparison of OK and KSMD estimations for soil texture, pH, M3-P, and M3-Ca, mean error values of KSMD are almost equal to those of OK.

This indicates KSMD estimation is as accurate as OK estimation.

Precisions of OK and KSMD estimations for soil texture, pH, M3-P, and M3-Ca

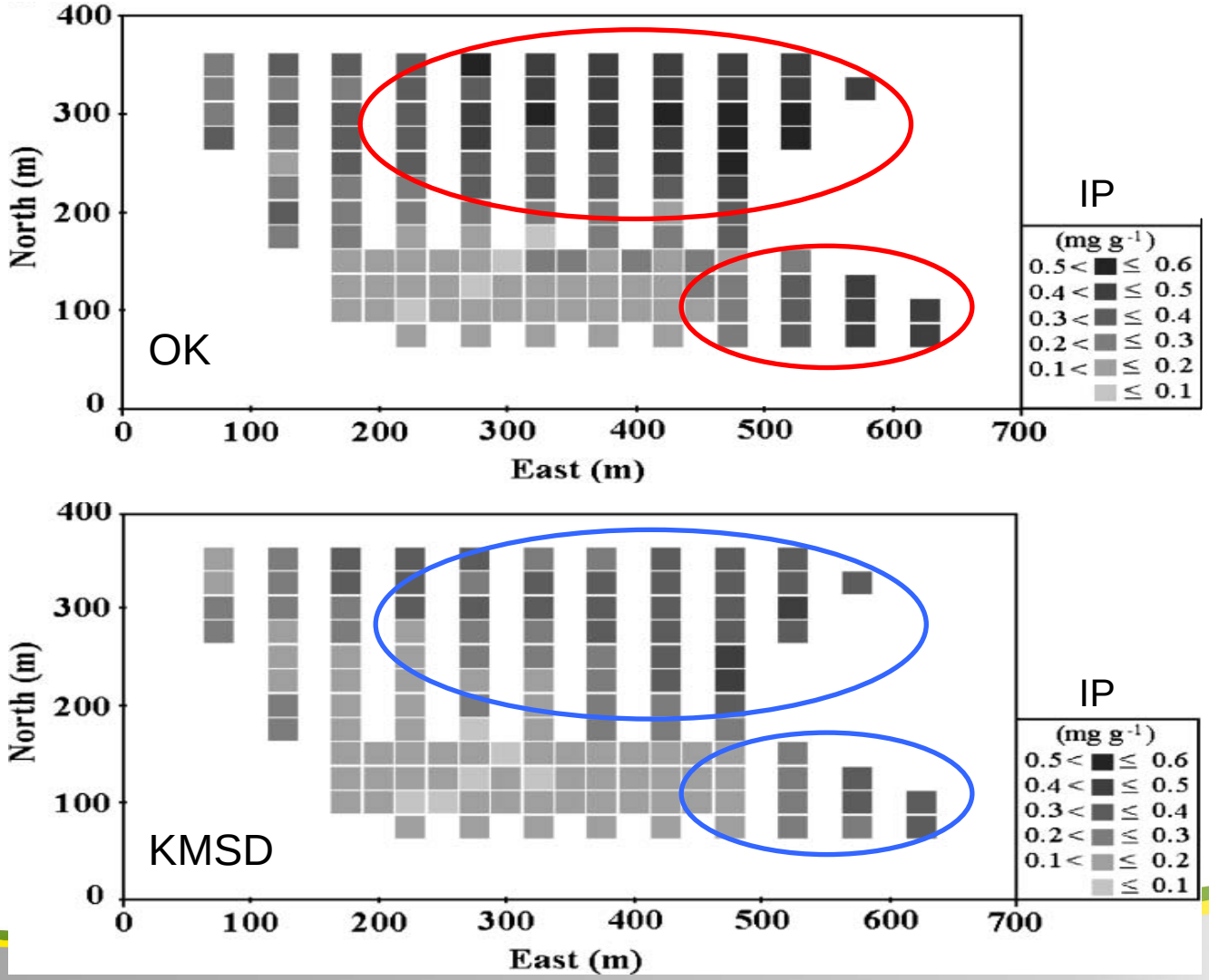


Comparison of precisions of OK and KSMD estimations

In comparison of OK and KSMD estimations for soil texture and M3-Ca, imprecision index values of KSMD are almost lower than those of OK.

This could be referred to soil type effects are significant on soil texture and M3-Ca and then the KSMD estimations presented higher precision than did the OK estimations.

Spatial assessment for estimation precisions of M3-Ca with OK and KSMD (n=127)



Spatial assessment for estimation precisions of M3-Ca with OK and KSMD (continued)

In spatial assessment for estimation precision of M3-Ca, the high-valued imprecision indices of KSMD are obviously lower than those of OK.

This indicates KSMD estimation is more precise than OK estimation.

Assessment for mapping quality of KSMD

Decrease of estimation imprecision of KSMD relative to OK (DIP):

$$DIP = \left[\frac{\sum_{i=1}^{n} IP^2(x_i)_{OK} - \sum_{i=1}^{n} IP^2(x_i)_{KMSD}}{\sum_{i=1}^{n} IP^2(x_i)_{OK}} \right] \times 100\%$$

n=127

	Sand	Silt	Clay	pH	M3-P	M3-Ca
DIP (%)	34	40	48	20	3	42

Assessment for mapping quality of KSMD (continued)

The decreases of estimation imprecision of KSMD relative to OK (DIP) for soil texture and M3-Ca are higher than 30%.

The mapping qualities with KSMD estimation for soil texture and M3-Ca were obviously improved.

Conclusions

- **KSMD is a derived kriging approach with categorical ancillary information.** The estimation accuracy of KSMD is similar to that of OK; However, compared with OK, the estimation precision is raised by using KSMD.
- The results suggested that **soil map–delineation could be used as auxiliary variables to improve the spatial interpolation of soil properties**.

Thank you for your attention