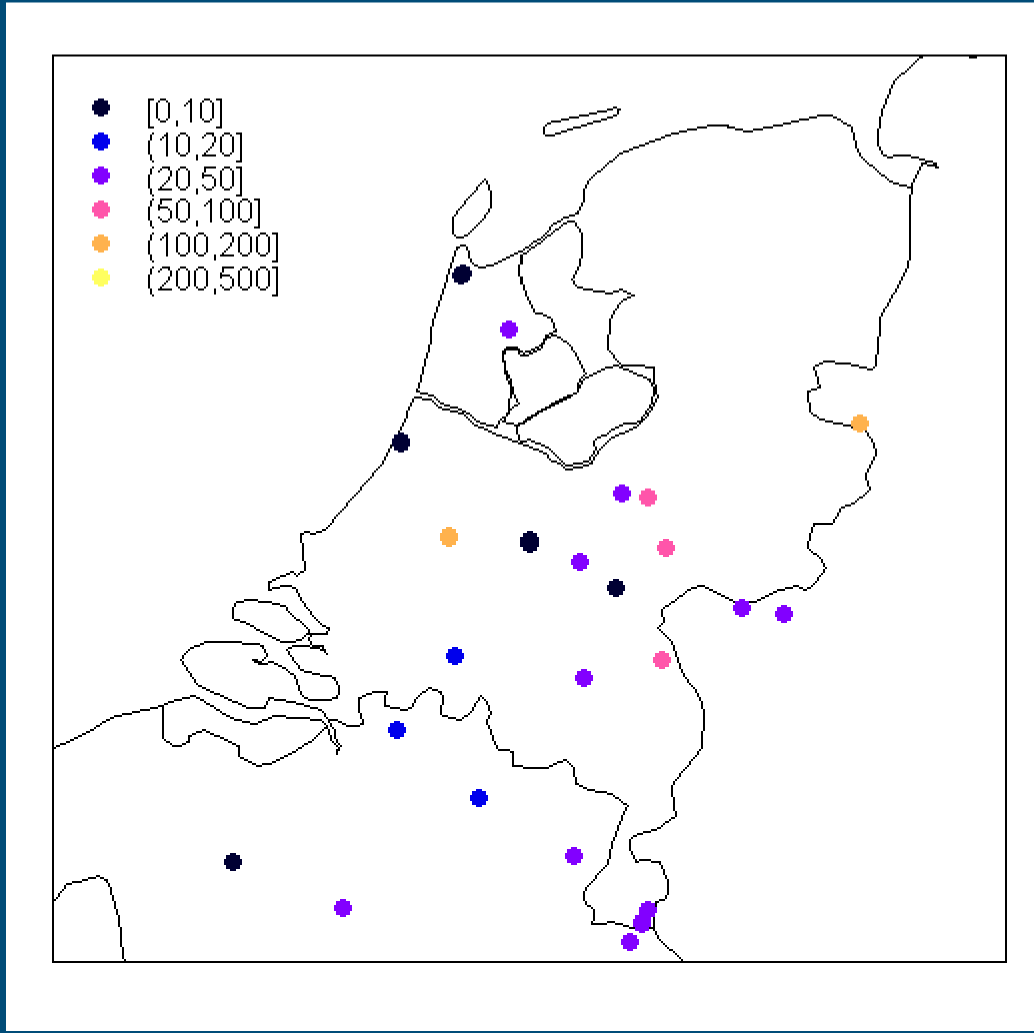
Incorporating process knowledge in spatial interpolation of environmental variables

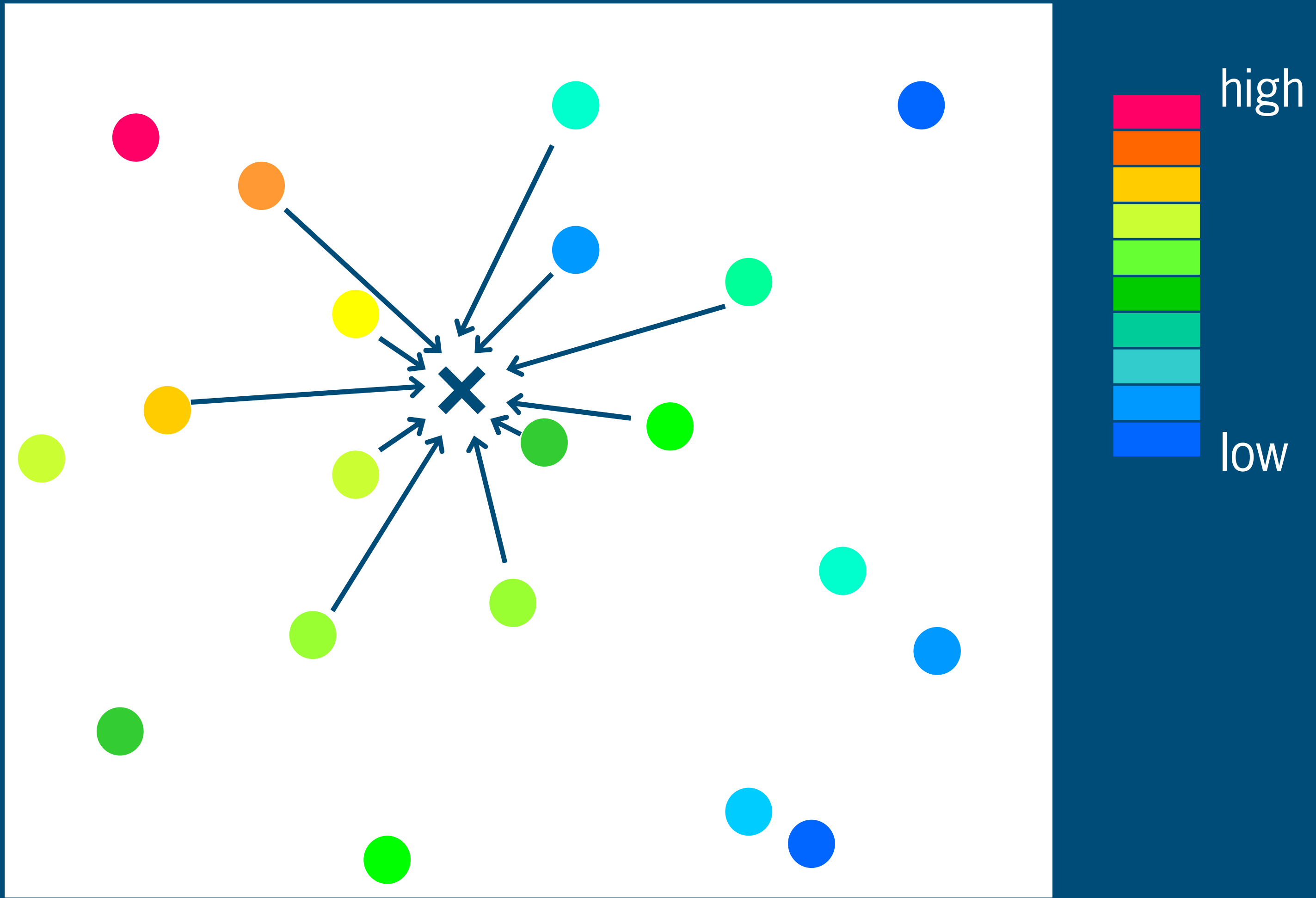
Gerard Heuvelink
Environmental Sciences Group
gerard.heuvelink@wur.nl

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Are 984 observations enough for geostatistical interpolation?



A typical geostatistical problem



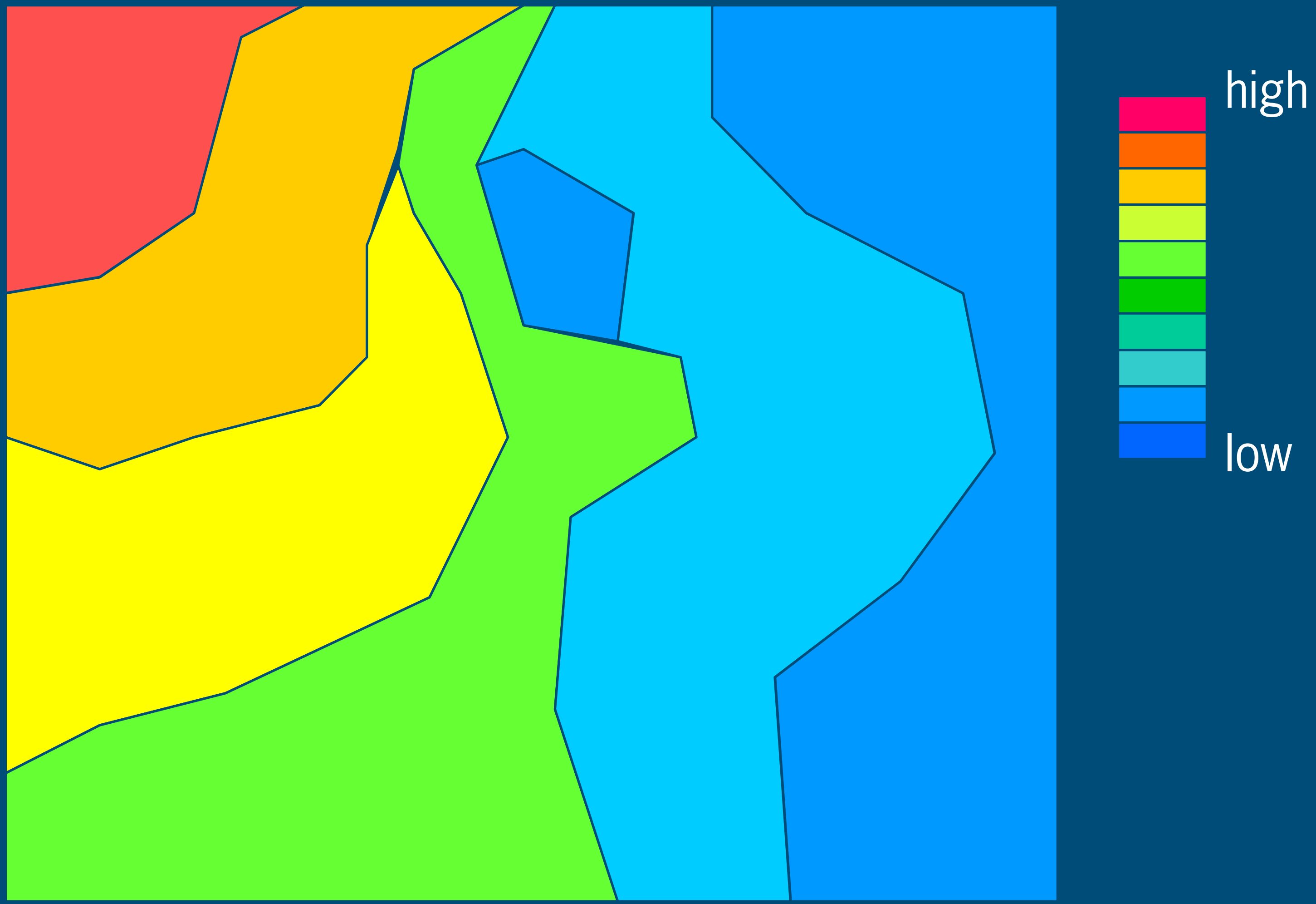
A typical geostatistical solution

Predict value of the spatial variable z at unobserved location x_0 from observations $z(x_i)$, $i=1\ldots n$, as follows:

$$\hat{z}(x_0) = \lambda_1 \cdot z(x_1) + \lambda_2 \cdot z(x_2) + ... + \lambda_n \cdot z(x_n)$$

Ordinary Kriging: derive weights λ_i from the spatial autocorrelation structure (semivariogram) of z, this yields the Best Linear Unbiased Predictor

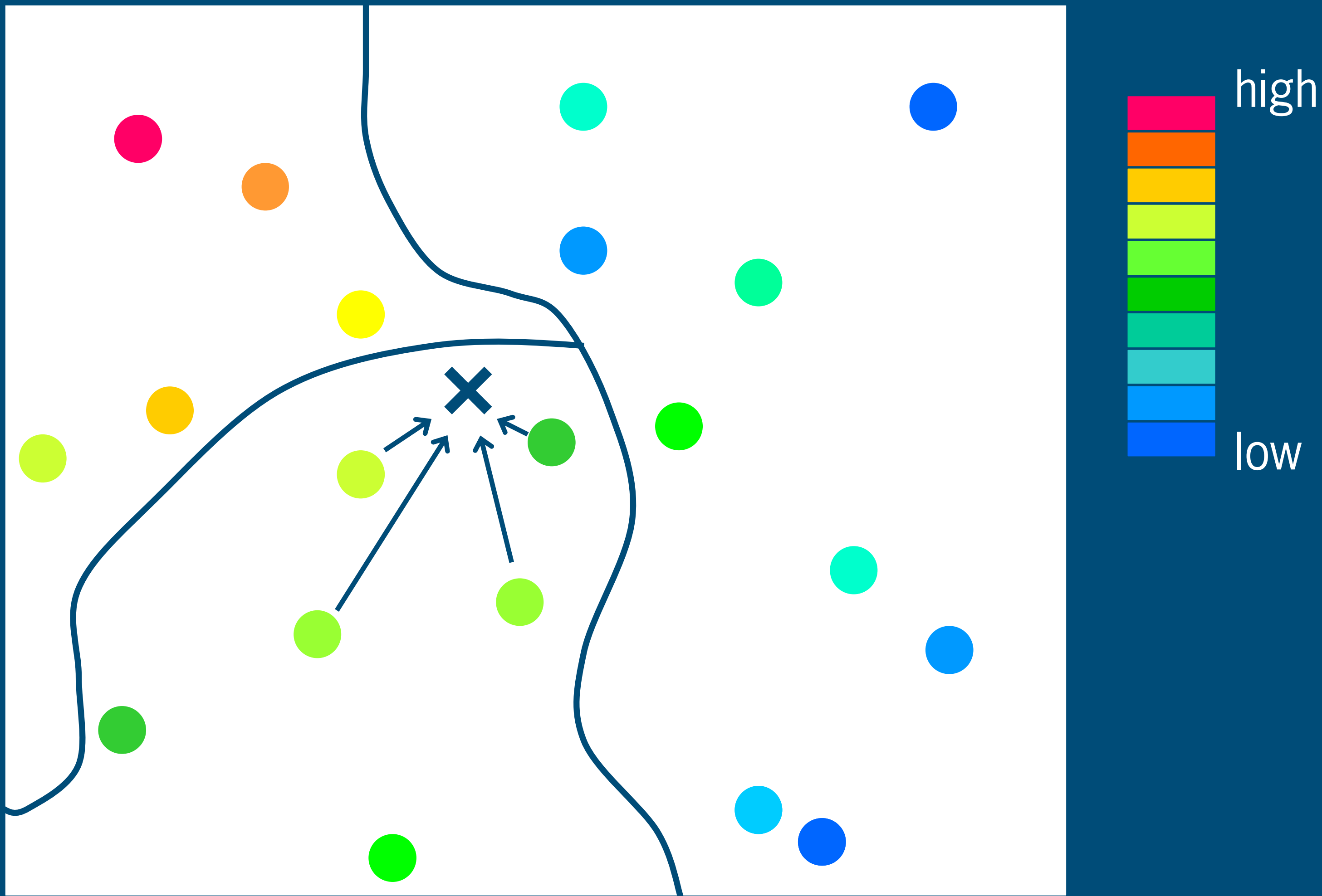
The result might look like this



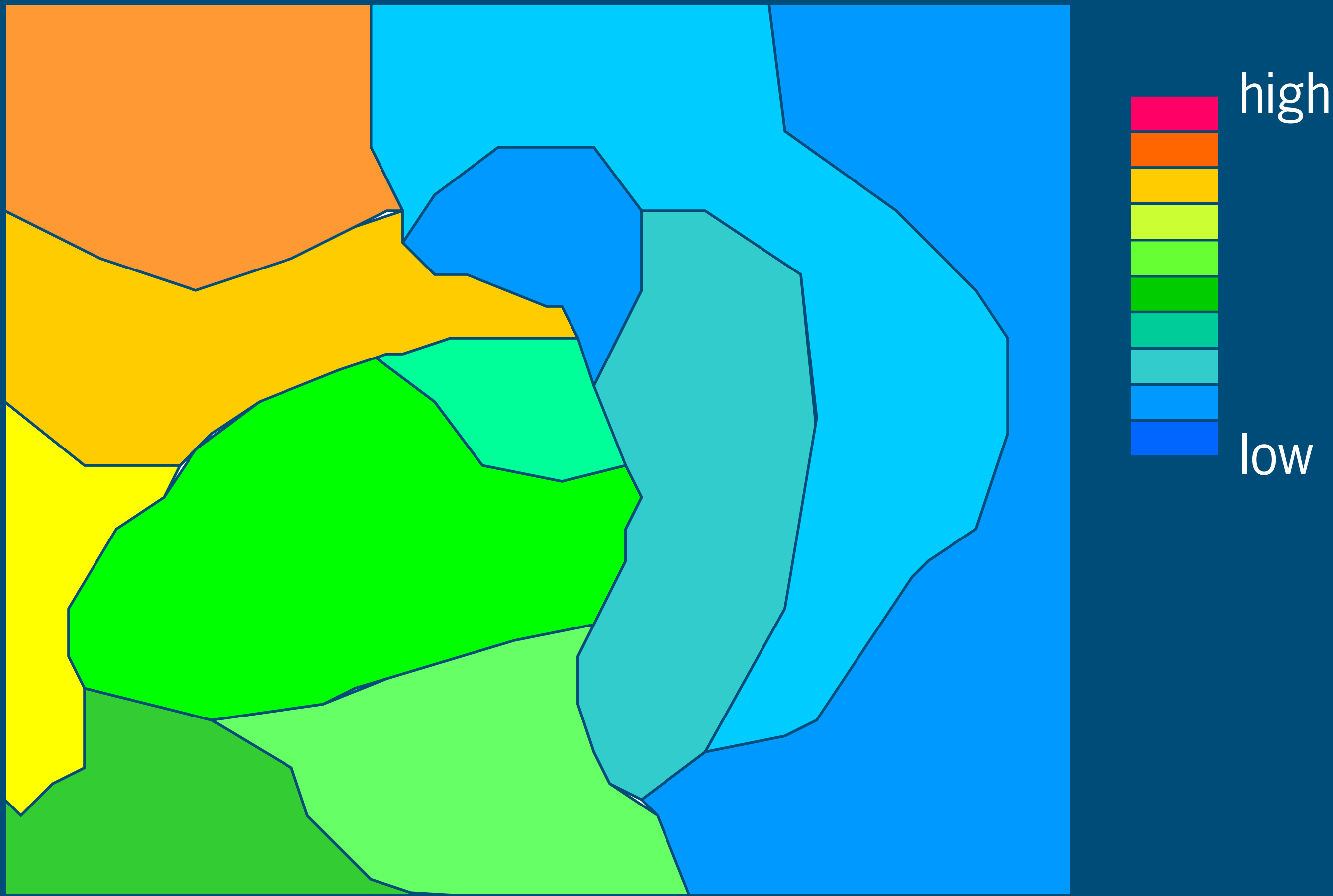
Problem solved?

- we have obtained a map of the spatial variable, which weighs the observations optimally
- however, ordinary kriging is entirely based on the observations and does not make use of any additional information (which is often available)
- perhaps we can do better by incorporating the additional information (explanatory data as well as knowledge about physical processes that caused the spatial variation)
- we will discuss three approaches to do so, starting simple but ending complicated

APPROACH 1. Stratified kriging



Stratified kriging preserves the boundaries between mapping units

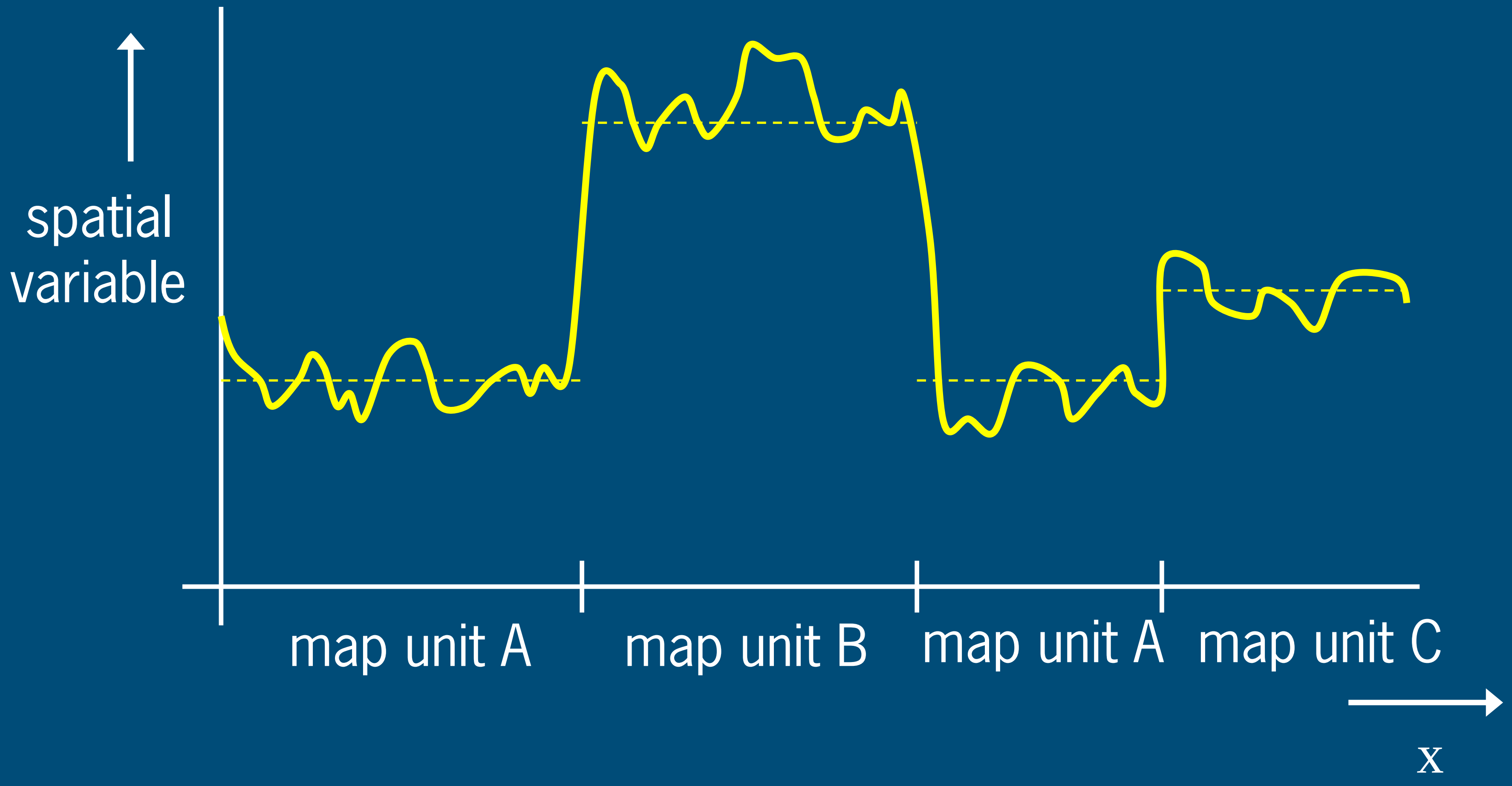


Statistical model underlying stratified kriging

$$z(x) = \underbrace{\text{mapping unit dependent mean at location } x}_{\text{(deterministic) trend, explanatory part}} + \underbrace{\text{deviation from the mean at } x}_{\text{(stochastic) residual, unexplanatory part}}$$

possibly spatially autocorrelated

Example realisation along a transect



WAGENINGEN UNIVERSITY
WAGENINGEN UR

APPROACH 2. Regression kriging

$z(x)$ = f(explanatory variables) + stochastic residual

possibly spatially autocorrelated

Example:

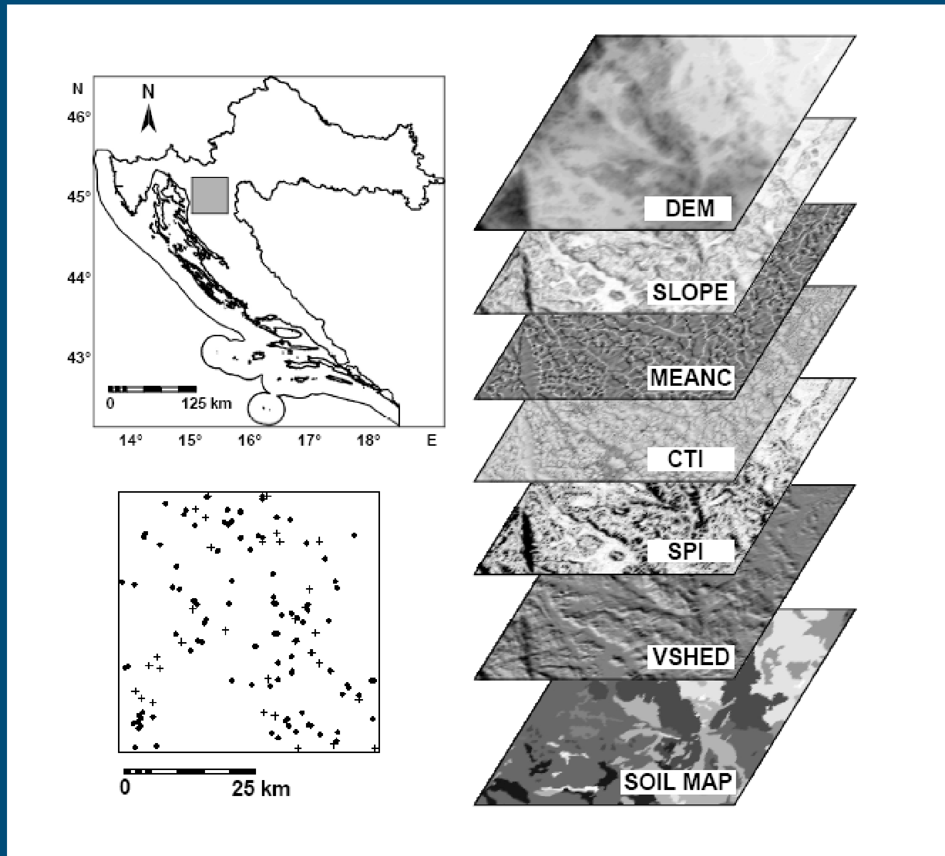
soil depth(x) = $\beta_0 + \beta_1 \cdot$elevation(x) + $\beta_2 \cdot$slope angle(x) + $\beta_3 \cdot$vegetation density(x) + $\beta_4 \cdot$upstream area(x) + residual (x)

Regression kriging algorithm

1. select explanatory variables and estimate regression coefficients using ordinary least squares
2. compute residuals (by subtracting the fitted trend from the observations) at observation locations and compute a semivariogram to quantify spatial correlation of the residual
3. apply the regression model at all unobserved locations (usually a grid)
4. krige the residuals
5. add up the results of steps 3 and 4

Better: integrate estimation of coefficients and kriging of residuals using weighted least squares and universal kriging

Example from Hengl et al. (*Geoderma* 120, pp. 75-93): predicting soil depth for a 50 × 50 km area in Croatia



Results using various interpolation methods

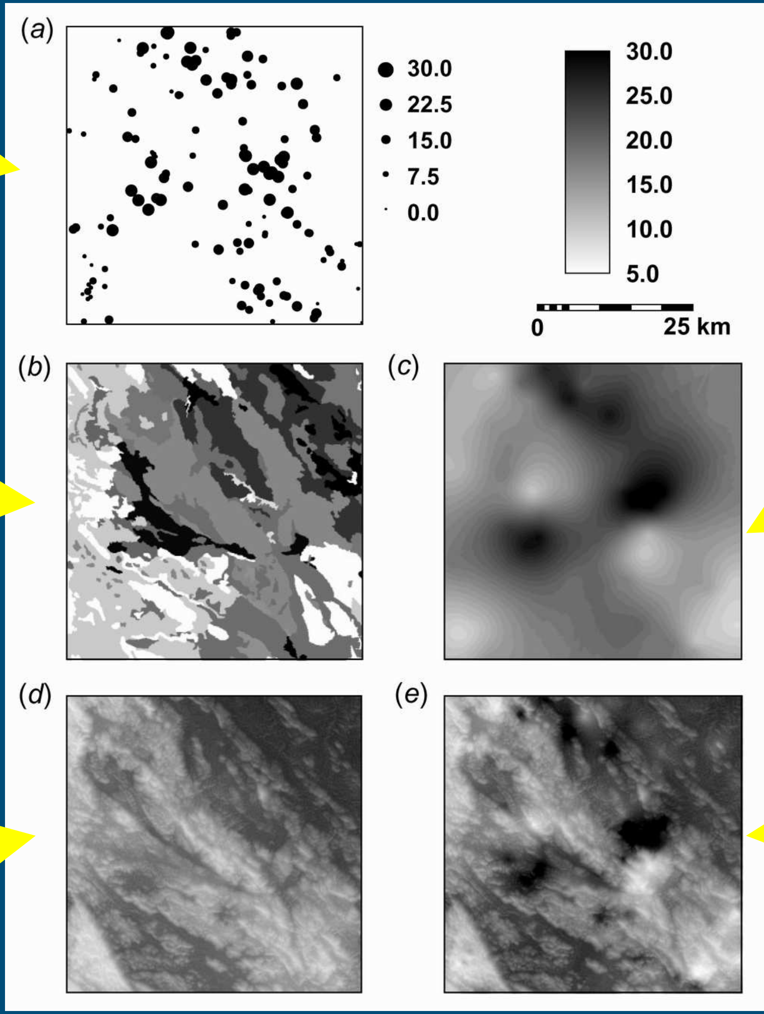
observations

soil map only predictor

ordinary kriging

regression only

regression kriging



Validation on 35 independent observations

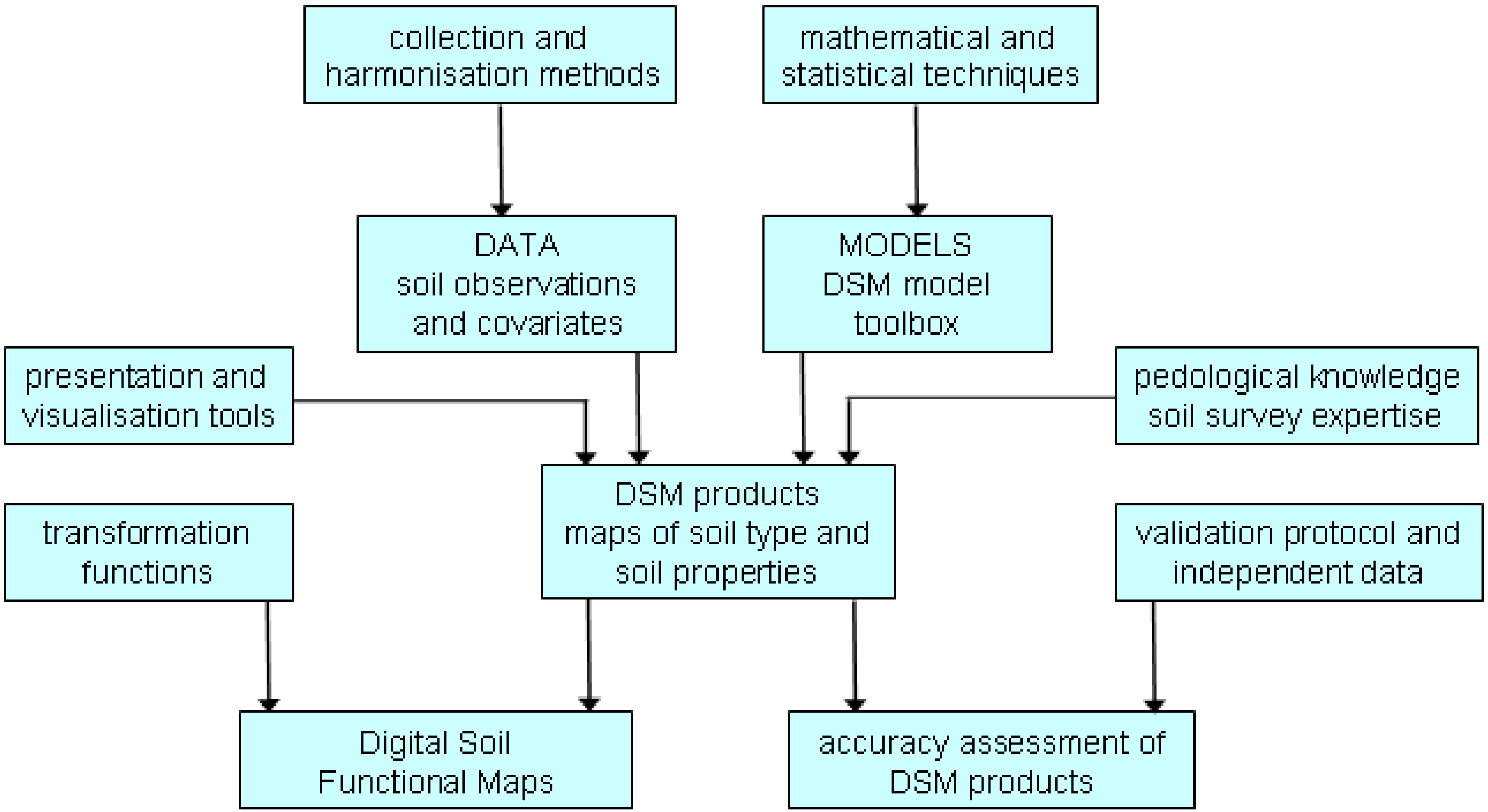
	Mean error [cm]	Root mean squared error [cm]
Soil map	1.42	9.1
Ordinary kriging	0.69	8.5
Multiple regression	1.69	8.8
Regression kriging	0.15	6.8

Regression kriging….

- is rapidly evolving because modern observation and GIS techniques yield high-quality explanatory variables at high resolutions
- incorporates process knowledge because it (presumably) uses explanatory variables that have a causal influence on the target variable
- is handicapped in the sense that the way in which explanatory variables appear in the trend is highly empirical, i.e. not reflecting the actual processes
- has given a boast to alternative ways of soil mapping, which has now entered the 'Digital Soil Mapping' era

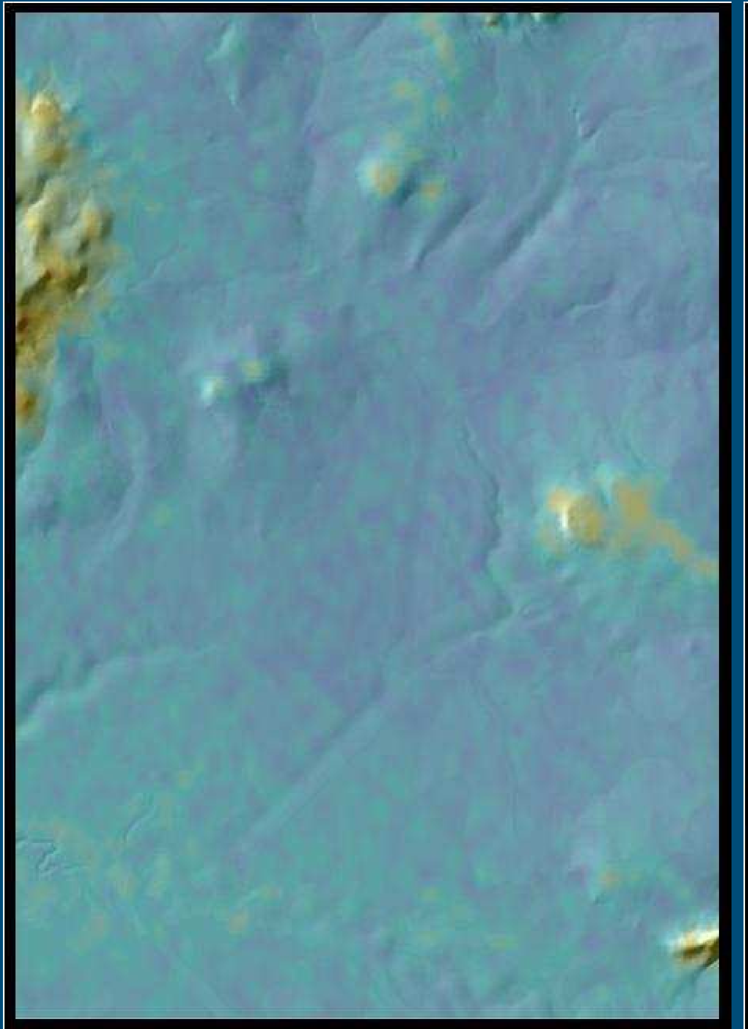

Global
Meeti
Home
News
Worksho
Committe
Program
Abstracts
Registrat
Accomm
Visitor In
Contacts
Print Flye
Portugue
Intranet
2008 usa
Digital Soil Mapping
3rd Global Workshop on Digital Soil Mapping
September 30 - October 3, 2008 • Utah State University • Logan, UT • USA
REGISTRATIO
Welco
"Di
Digital So
referenced
laboratory
quantitative
The Interna
(http://www
Soil Mappin
Soil Scienc
The theme
Mapping: A
2006) in Ri
with Sparse
GlobalSoilMap.net
HOME
WORLD SOIL MAPS
PROPERTIES
SOFTWARE
PUBLICATIONS
CONSORTIUM
LINKS
CONTACTS
News
Download the press release
Download the brochure
GlobalSoilMap.net in the news
The project was officially launched 17th February, New York, USA
presentations programme
speaker biographies outcome
The African part of GlobalSoilMap.net was launched on 13th January 2009 in Nairobi. Read here the press coverage www.africasoils.net
"Let there be no mistake about the significance of this wonderful project"
Kofi Annan
"Soil mapping is one of the pillars to the challenge of sustainable development"
Jeffrey Sachs
17th February 2009
There is a need for accurate, up-to-date and spatially referenced soil information. This need has been expressed by the modelling community, land users, and policy and decision makers. This need coincides with a enormous leap in technologies that allow for accurately collecting and predicting soil properties.
We have formed a consortium that aims to make a new digital soil map of the world using state-of-the-art and emerging technologies for soil mapping and predicting soil properties at fine resolution. This new global soil map will be supplemented by interpretation and functionality options that aim to assist better decisions in a range of global issues like food production and hunger eradication, climate change, and environmental degradation. This is an initiative of the Digital Soil Mapping Working Group of the International Union of Soil Sciences IUSS
In November 2008, an $18 million grant has been obtained from the Bill & Melinda Gates foundation and the Alliance for a Green Revolution in Africa

Digital Soil Mapper



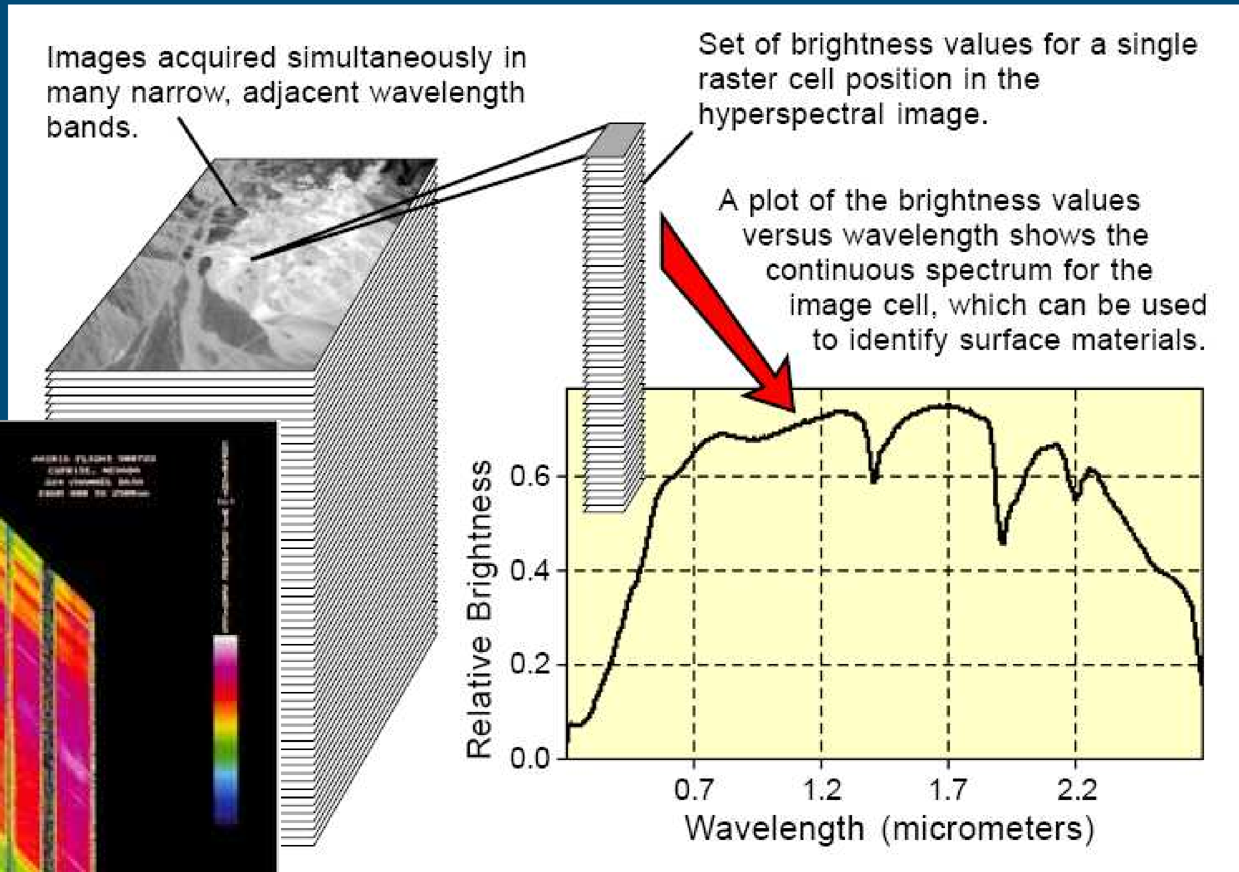
DEM resolution ever increasing

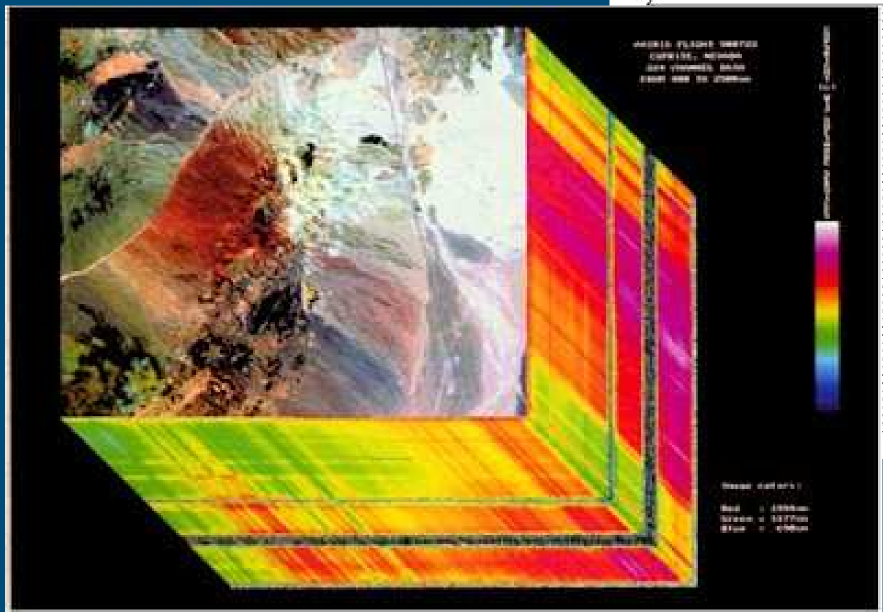
DEM source: 1:50 000 contours
Radiometrics: 200 m line spacing



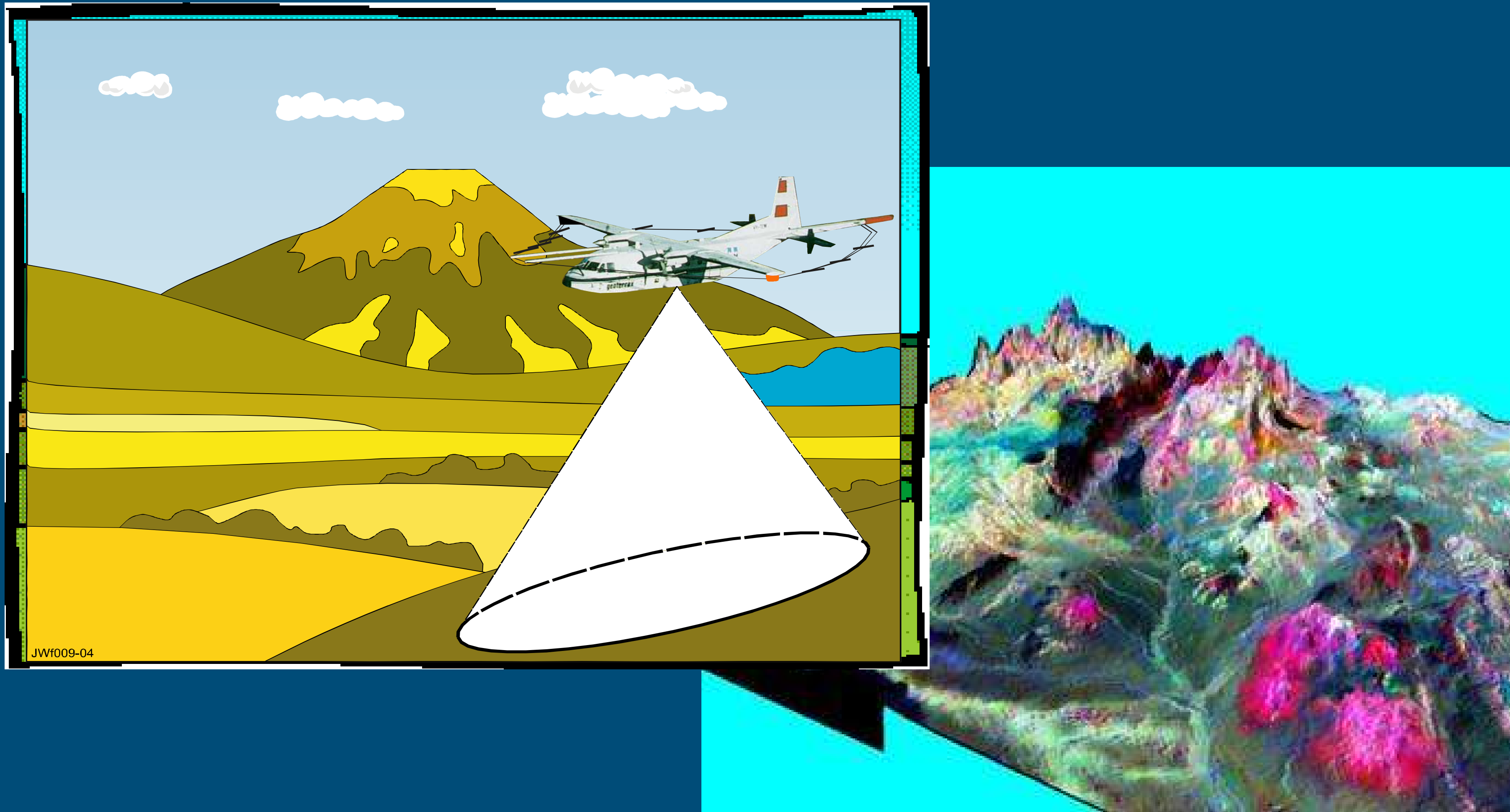
DEM source: LASER Altimetry
Radiometrics: 50 m line spacing

High resolution imagery, also in feature space

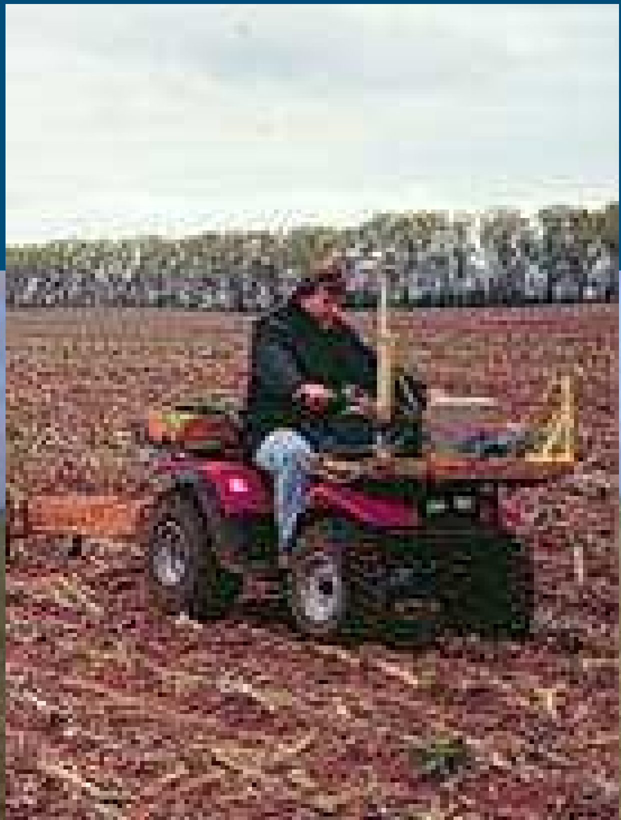


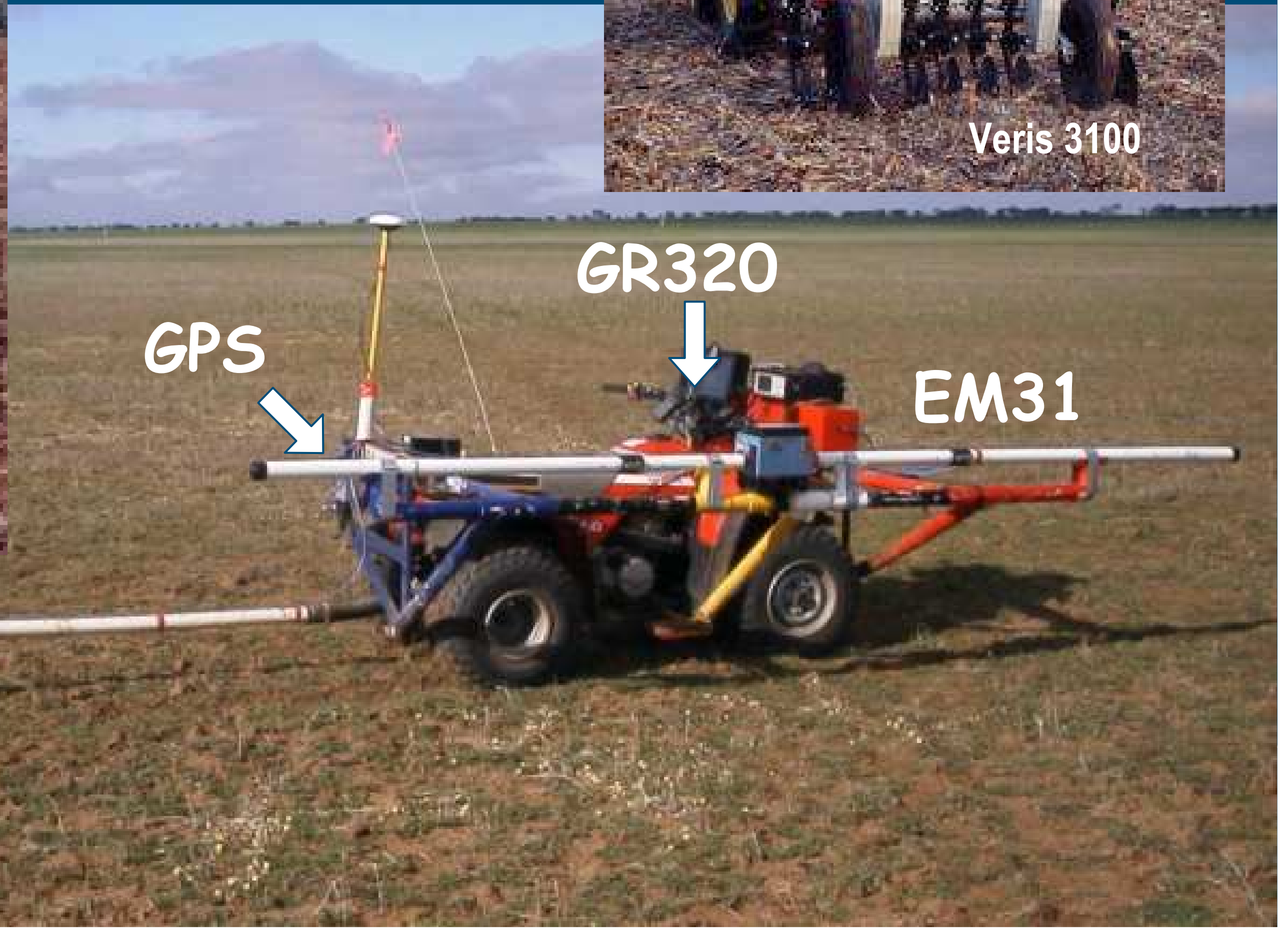


Airborne geochemistry



New ground surveying techniques





Regression kriging....

- is handicapped in the sense that the way in which explanatory variables appear in the trend is highly empirical, i.e. not reflecting the actual processes

APPROACH 3. Space-time Kalman filtering

To do better justice to process knowledge we must take a dynamic approach

state of system (t) = f(state of system(t–1), external forces) + residual

proce model

spatially distributed!!

no matter how hard we try, the model will always differ from reality

State-space approach has two main equations

State equation (assume linear model):

$$Z(t+1) = A(t) \cdot Z(t) + B(t) \cdot U(t) + \varepsilon(t), \qquad t \geq 0$$

system state

system noise
= model error

external forcing

Measurement equation:

$$Y(t) = C(t) \cdot Z(t) + \eta(t), \qquad t \geq 0$$

measurement

measurement error

Kalman filter algorithm: combine process knowledge with information in measurements

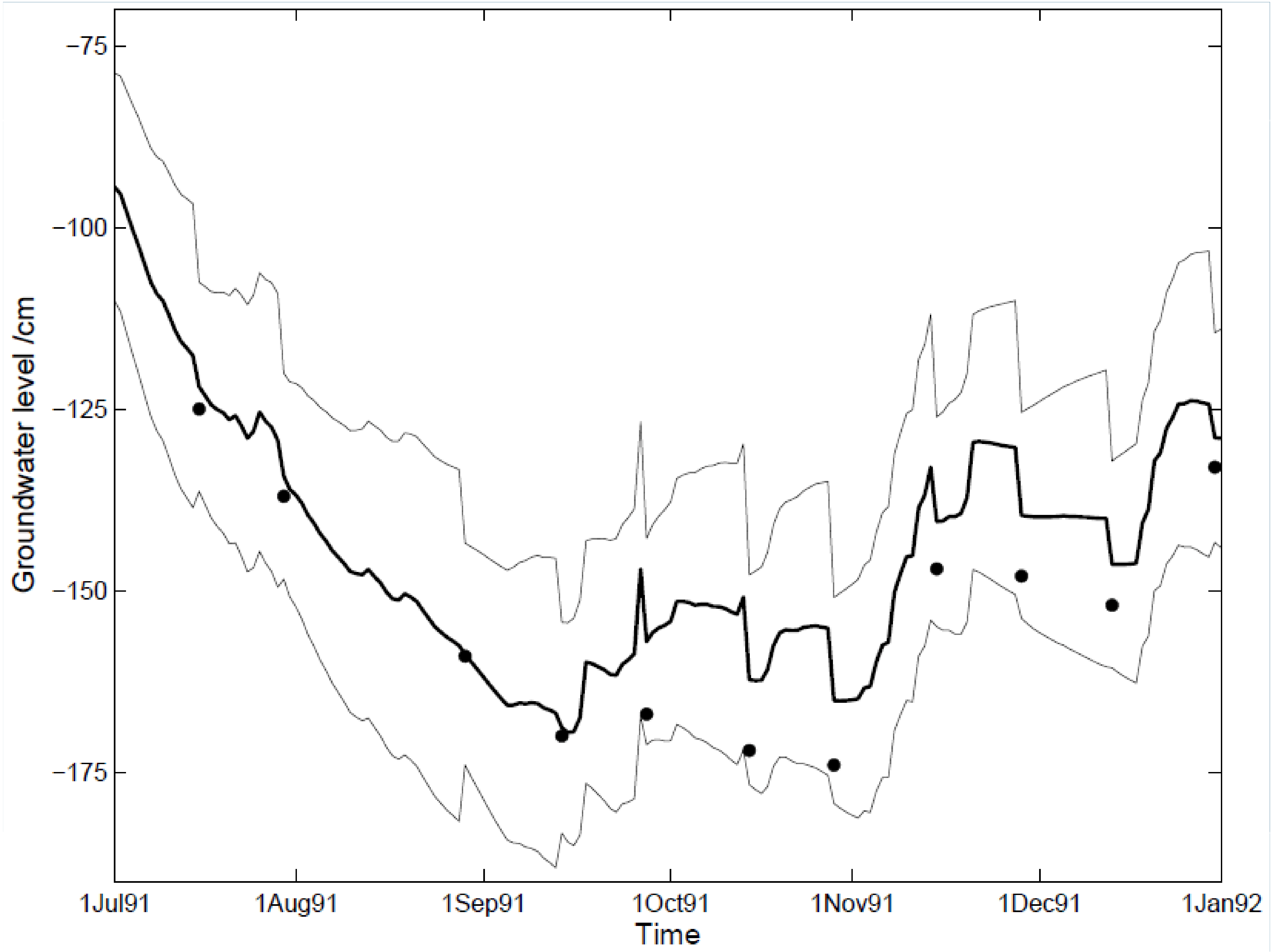
Starting from state Z_0 at $t=0$, we have a time update:

$$\hat{Z}^{-}(t+1) = A(t) \cdot \hat{Z}^{+}(t) + B(t) \cdot U(t)$$

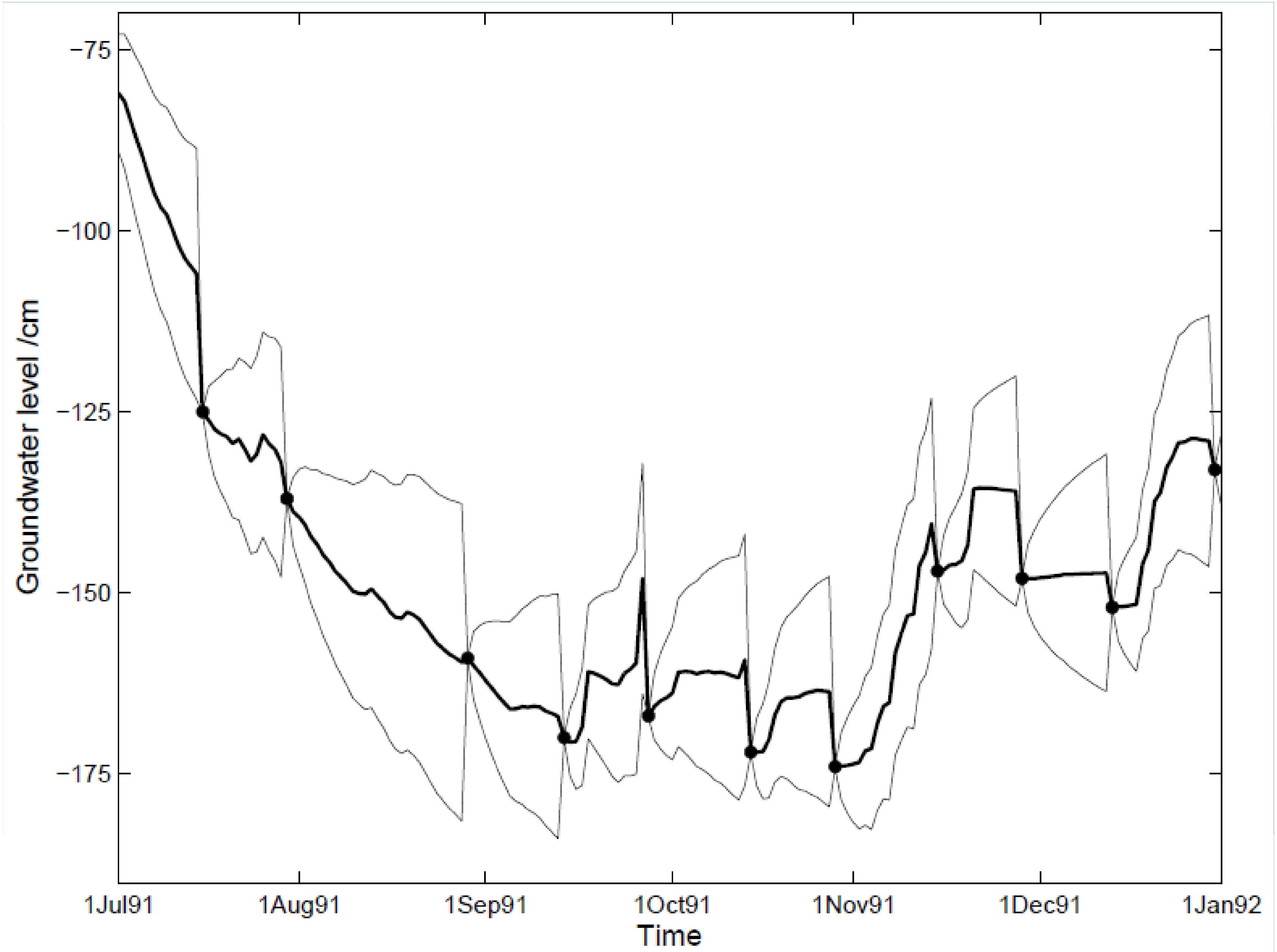
and a measurement update:

$$\hat{Z}^{+}(t+1) = \hat{Z}^{-}(t+1) + K(t+1) \cdot (Y(t+1) - C(t+1) \cdot \hat{Z}^{-}(t+1))$$

where $K(t+1)$ is the Kalman Gain, which determines how much weight the measurement gets to correct the state estimate. It can be computed alongside the updates, and so can the associated variances



−75
−100
−125
−150
−175
Groundwater level /cm
1Jul91
1Aug91
1Sep91
1Oct91
1Nov91
1Dec91
1Jan92
Time

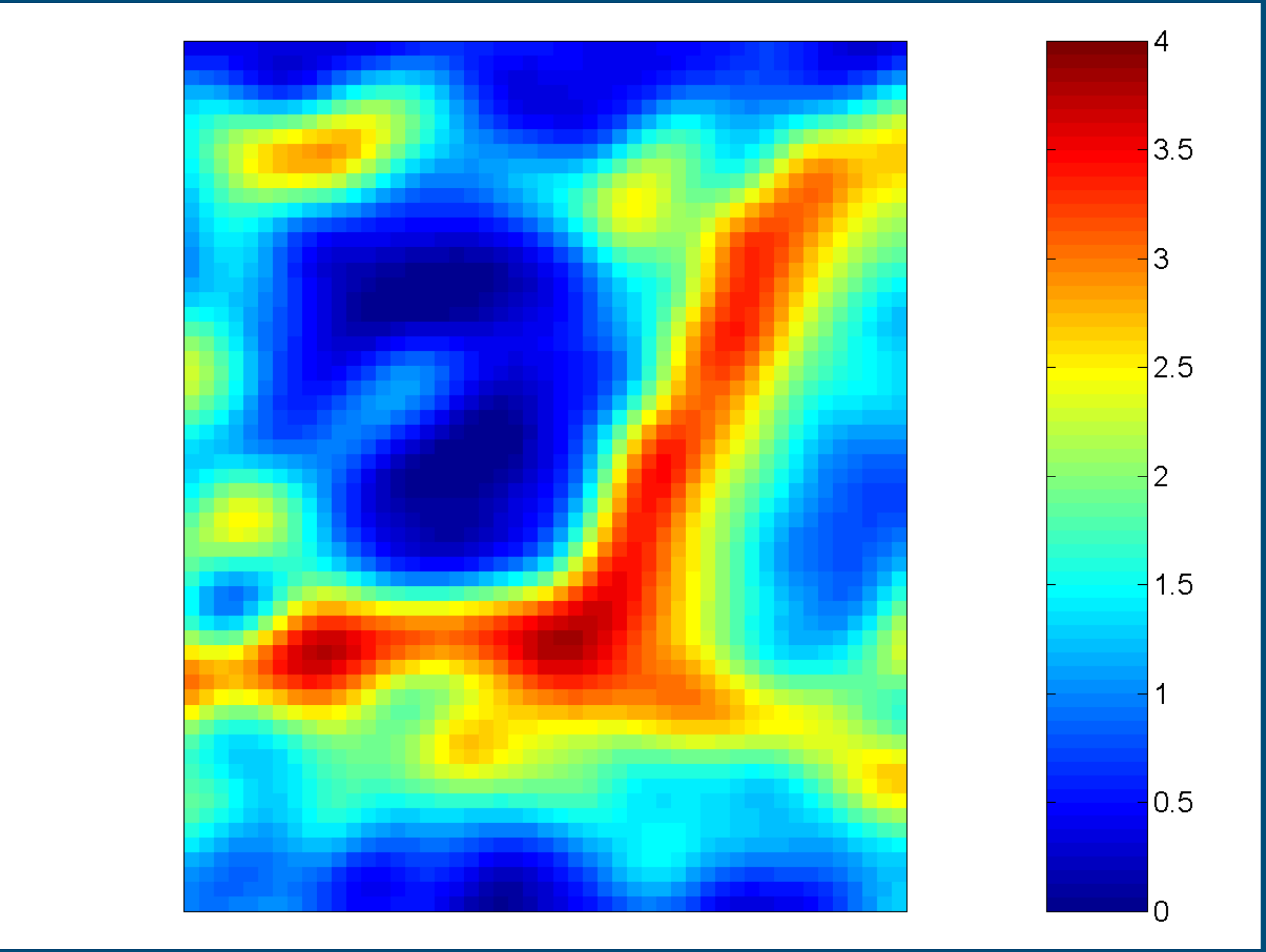


−75
−100
−125
−150
−175
Groundwater level /cm
1Jul91
1Aug91
1Sep91
1Oct91
1Nov91
1Dec91
1Jan92
Time

Application of the space-time Kalman filter to mapping soil redistribution in the Hepburn research site

- Hepburn site about seven hectares in size, located in southern Saskatchewan (Canada)
- gentle slopes, maximum difference in elevation three metres
- agricultural landuse (crop-fallow production system)
- tillage erosion main cause of soil redistribution
- amount of soil flux per tillage event has a linear relationship with slope angle (‘realistic’ assumption)

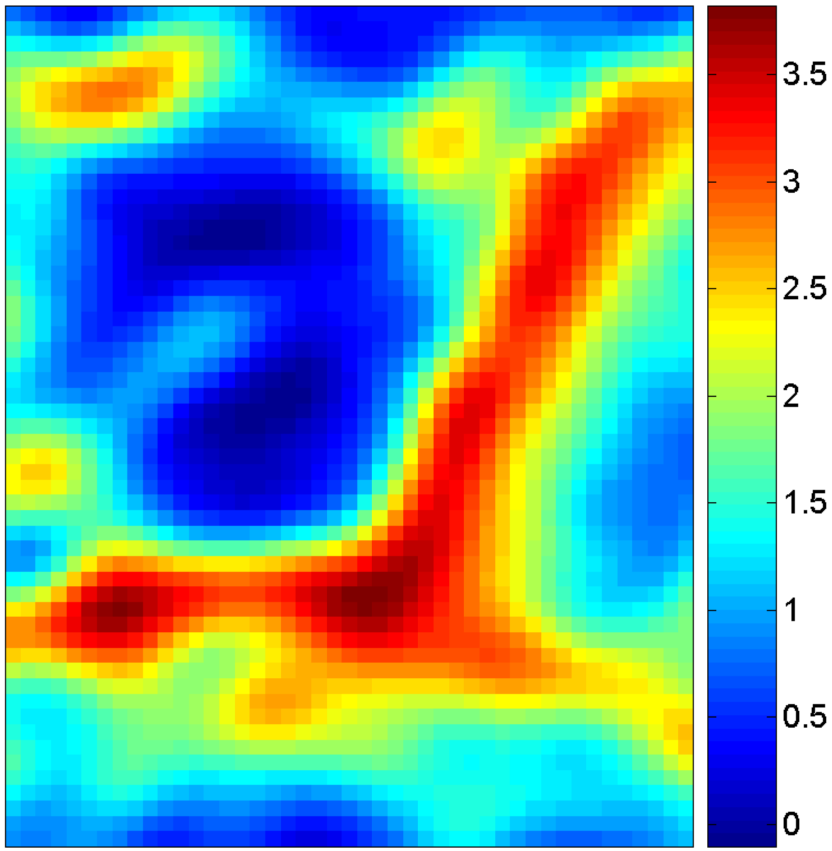
Initial DEM of Hepburn site



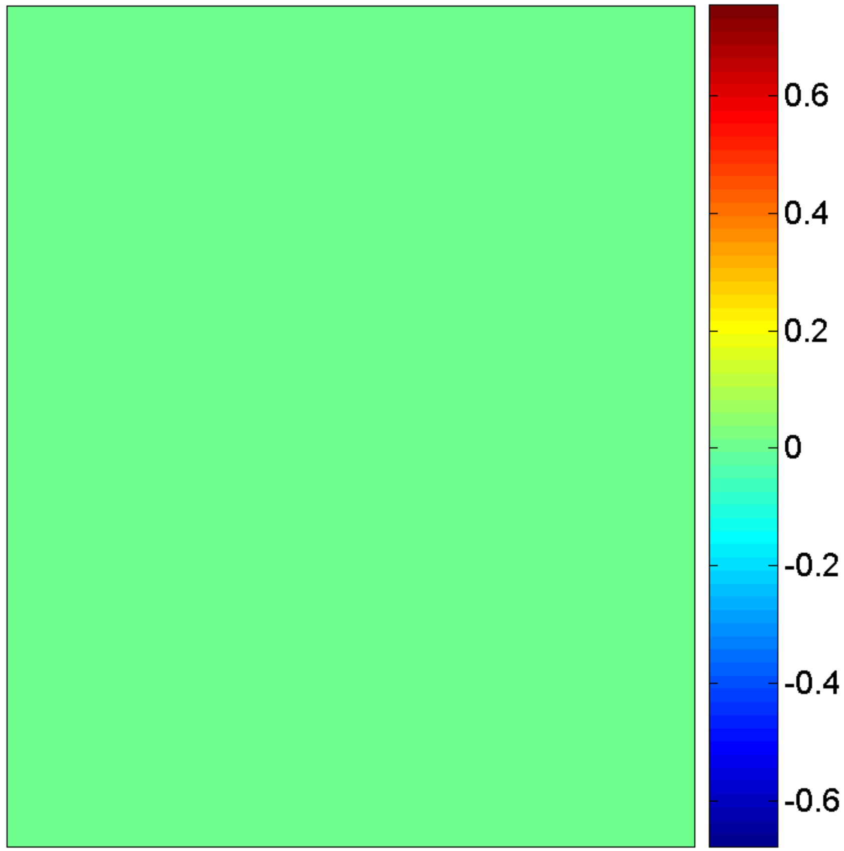
WAGENINGEN UNIVERSITY
WAGENINGEN UR

Evolution of soil redistribution over 37 years using process model only

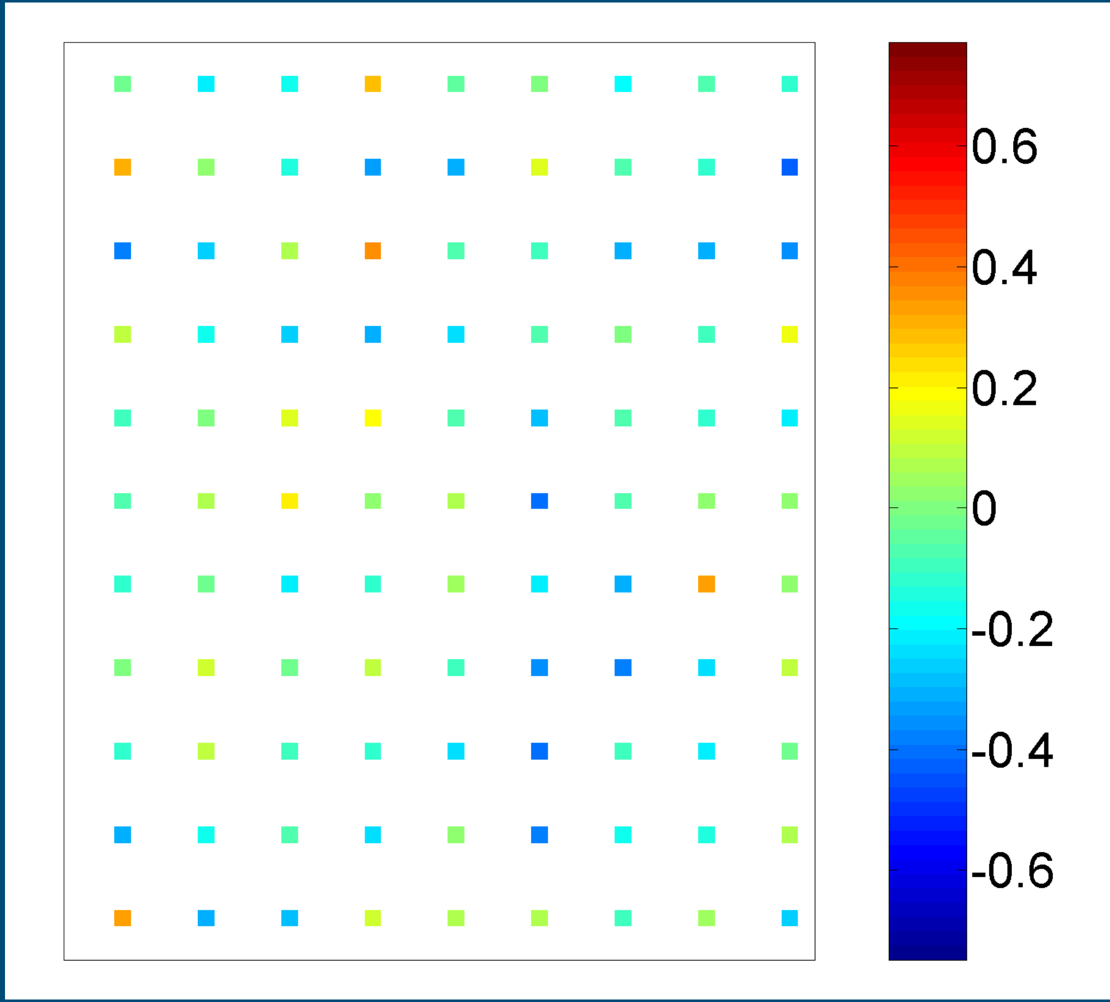
elevation



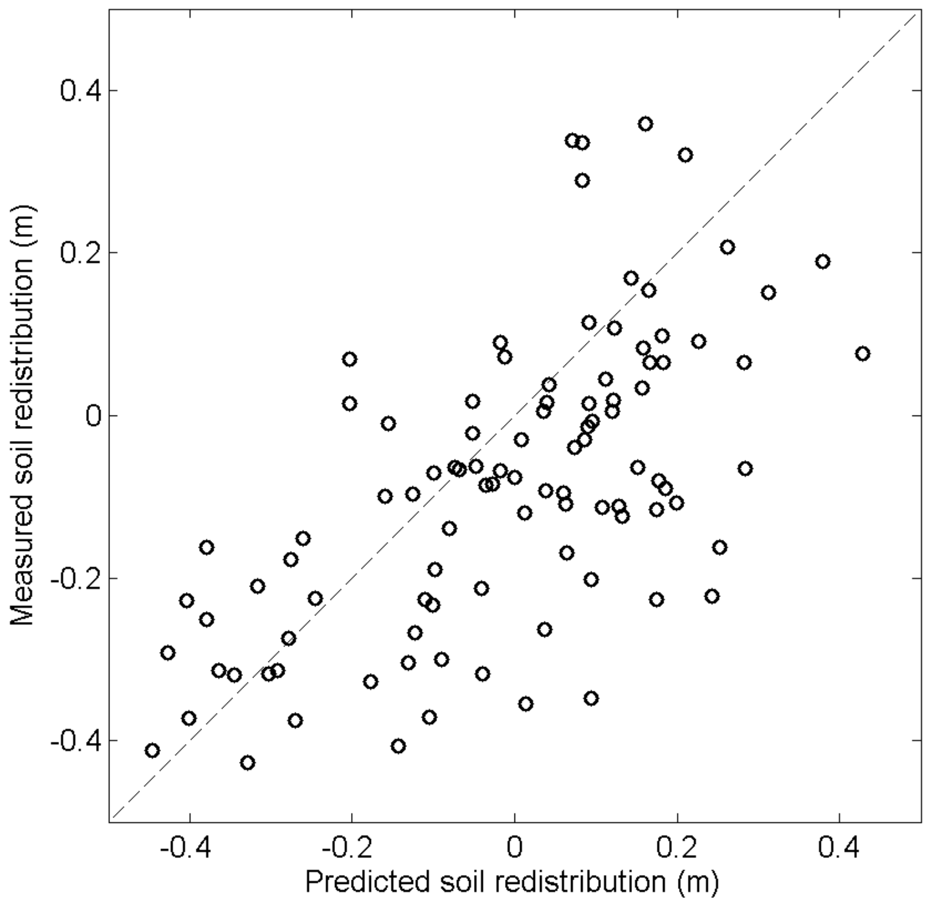
erosion/sedimentation



99 grid measurements of cumulative soil redistribution (sum over 37 years)

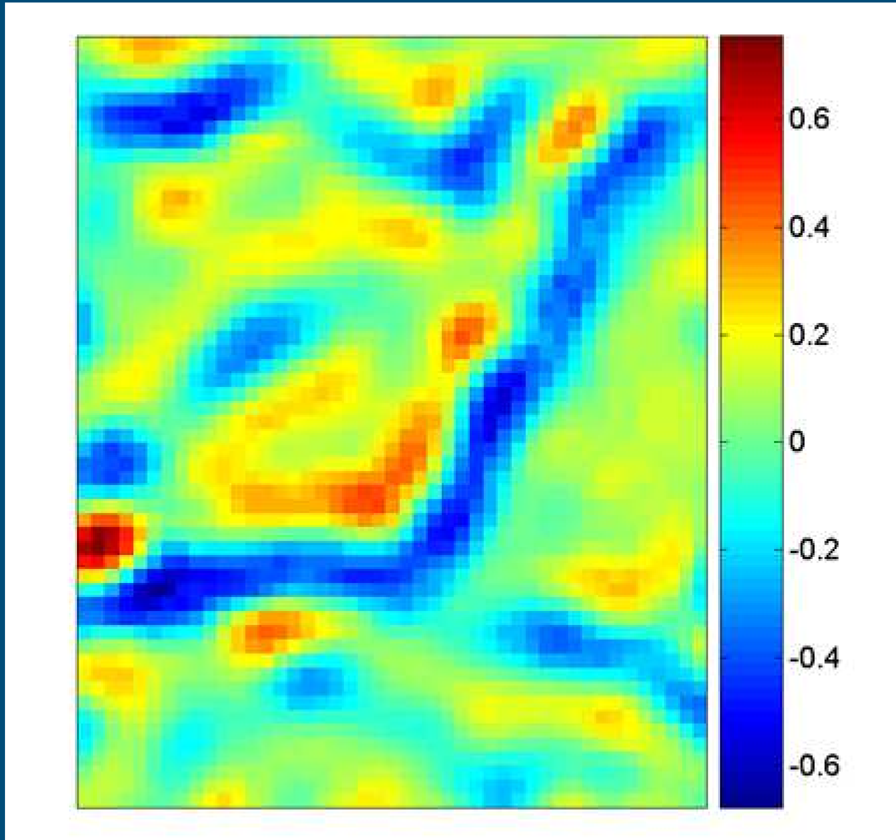


Scatter plot of measurements against process model predictions

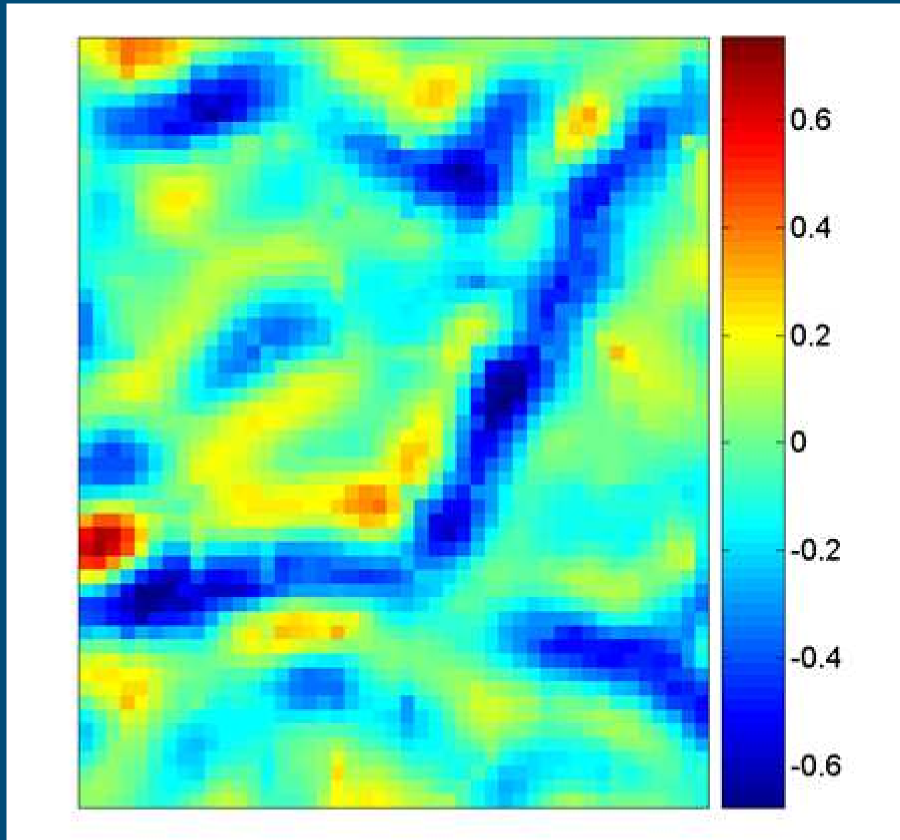


The space-time Kalman filter adjusts the predicted soil redistribution to the measurements

Before measurement update:

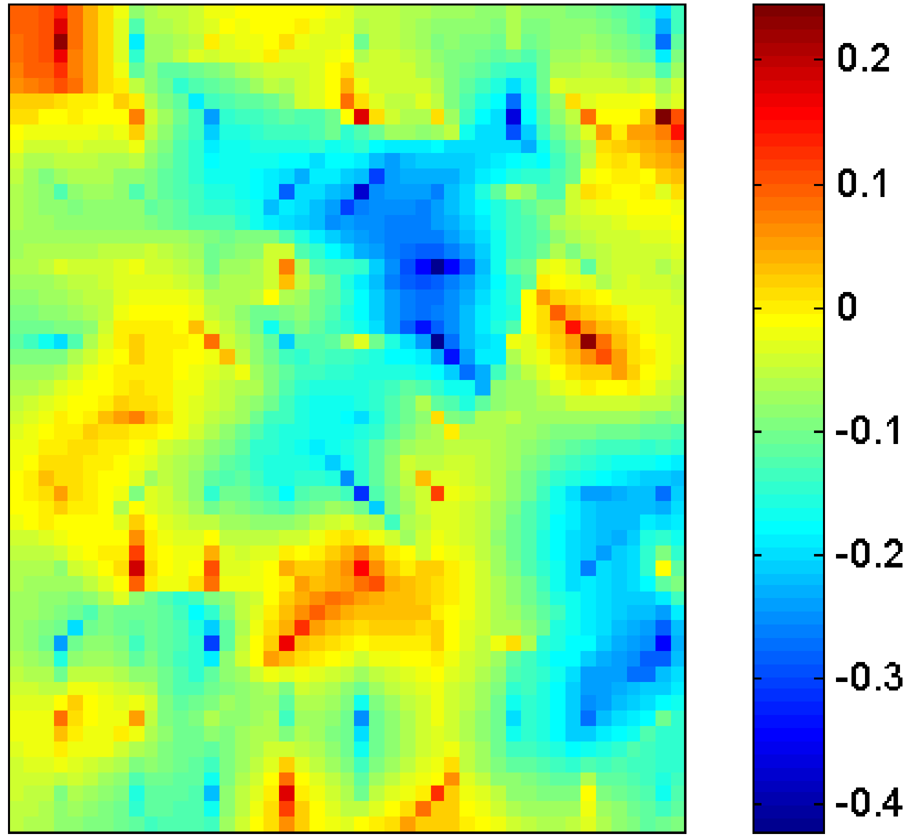


After measurement update:

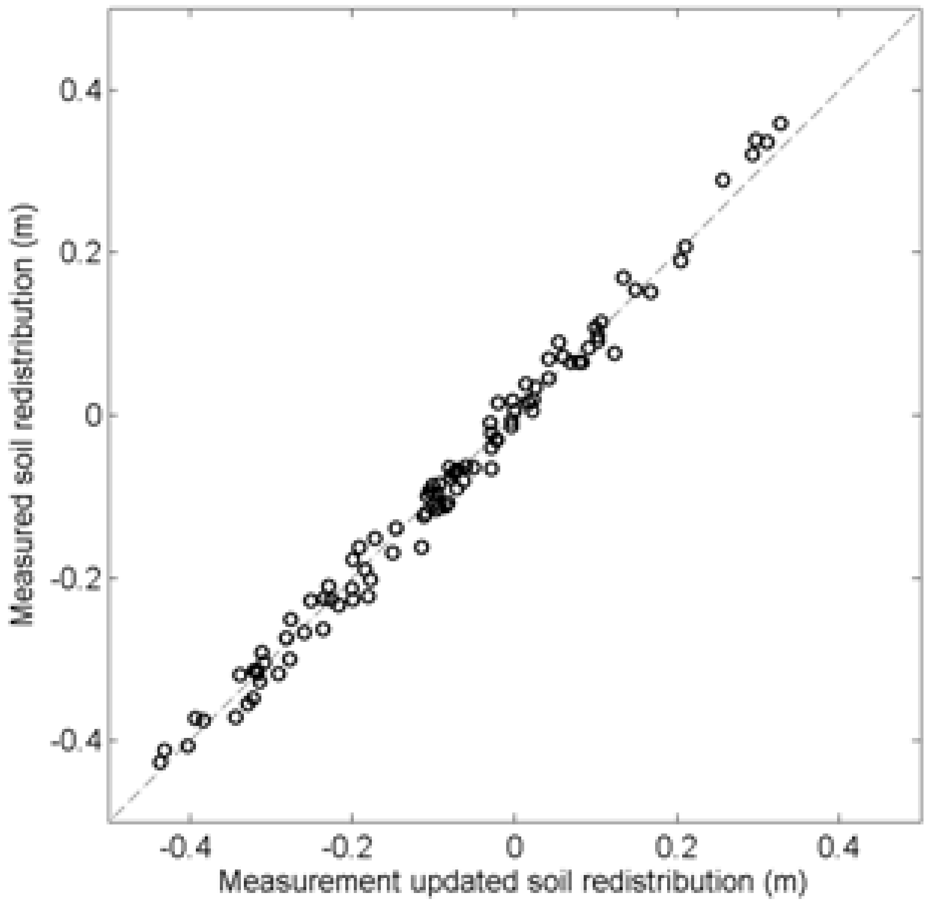


Marked adjustment, particularly along transportation routes near measurement locations

Effect of measurement update ('interpolated residual')



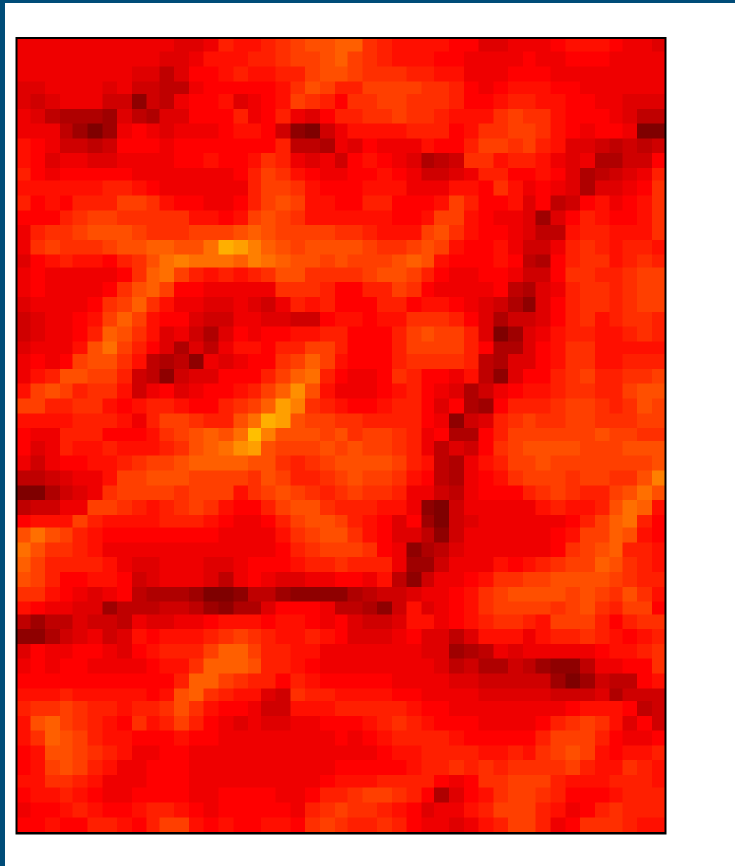
Scatter plot of measurements against updated model predictions

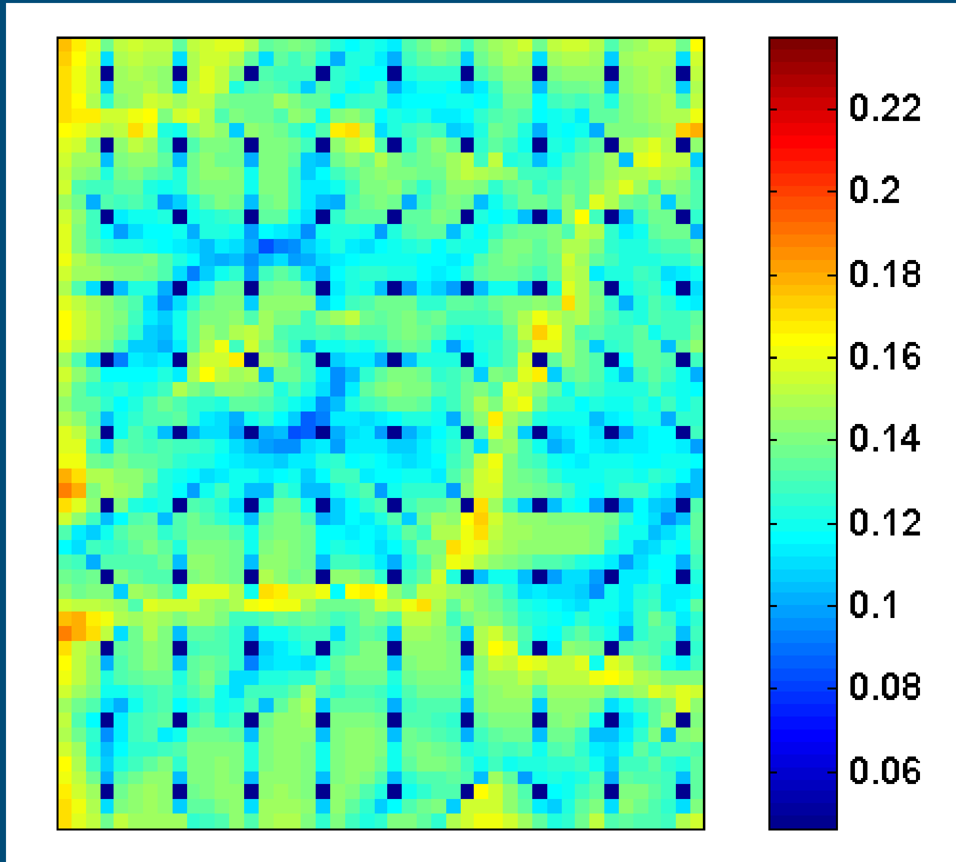


Measurement update also reduces uncertainty

Before measurement update

After measurement update





Summary and Conclusions

- There is much that can be gained by including process knowledge in spatial interpolation
- Model of spatial variation underlying spatial interpolation:

  variable = trend/explanatory part + stochastic residual
- Ordinary kriging focuses entirely on the residual and exploits its spatial autocorrelation
- Regression kriging pays more attention to the explanatory part
- Space-time Kalman filter represents real-world processes more realistically by taking a dynamic approach, while taking process model error into account and using measurements to correct the model predictions
- The advantage of exploiting process knowledge is not only that we (potentially) get more accurate maps, but also that we get a better understanding of how the real world works: that is what science is all about, is it not?

Thank you