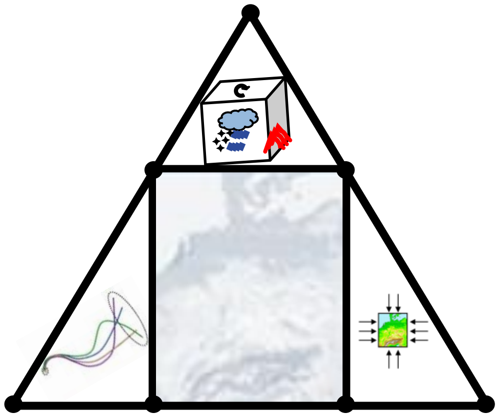
Stochastic modelling of the model error



Recent developments and open questions

$$\frac{\partial \psi}{\partial t} = \left[\frac{\partial \psi}{\partial t}\right]_{\mathrm{det}} + \eta(t)$$

$$\frac{\partial \eta}{\partial t} = \ldots\ldots\ \sigma\xi(t)$$

ICON-EPS

M. Sprengel, T. Heppelmann, E. Machulskaya, C. Gebhardt

Outline

- reminder: Ekaterina's scheme
- alternative: Stochastic Pattern Generator (Tsyrulnikov & Gayfulin)
- comparison w.r.t. theoretical aspects and practical implementation
- parameter estimation for the error tendency model
  (COSMO-D2 and ICON)
- next steps and open questions

EM-scheme *model for the model error (E. Machulskaya)*

$$\frac{\partial \psi}{\partial t} = \left[\frac{\partial \psi}{\partial t}\right]_{\mathrm{det}} + \eta(t) \qquad \frac{\partial \eta}{\partial t} = -\gamma\eta + \gamma\nabla(\lambda^2\nabla\eta) + \sigma\xi(t)$$

ψ : prognostic variables (T, QV, U, V)

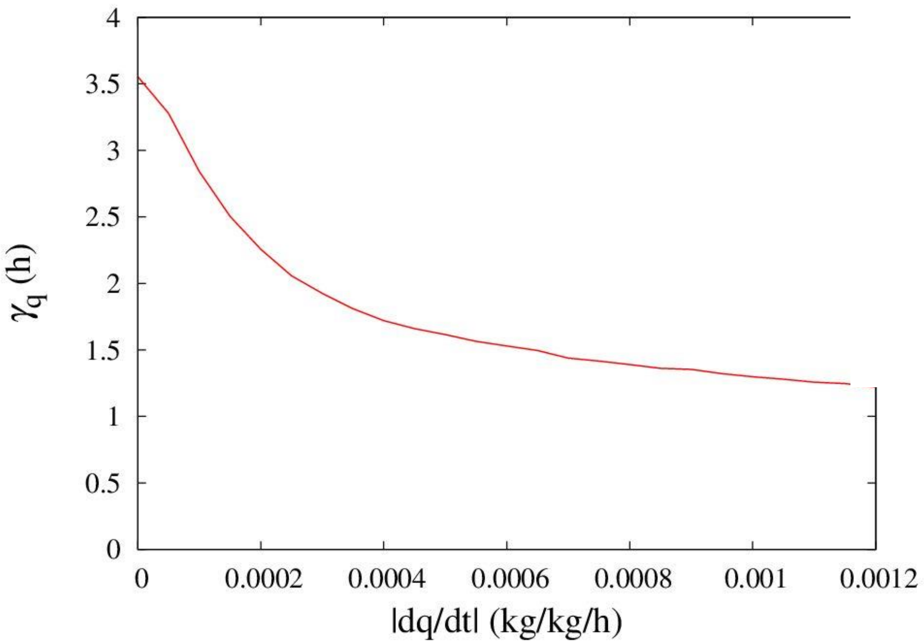
$\eta(t)$: noise field / model error, correlated in time and space

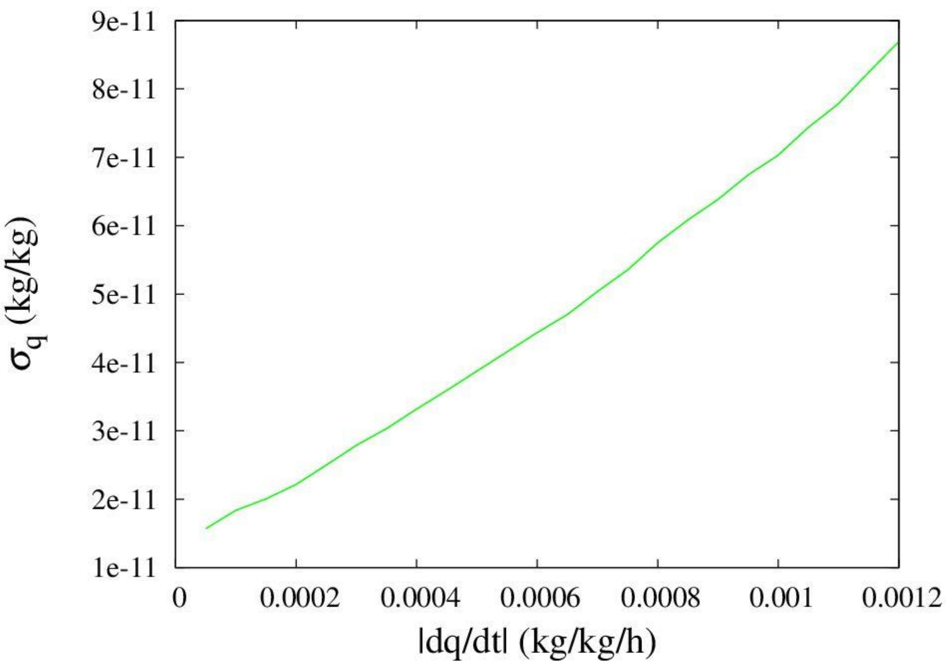
$\xi(t)$: Gaussian noise

σ, γ, λ: standard deviation and spatial and temporal correlation

> γ, λ and σ are weather-dependent and are derived from past data
>
> Potential predictors are $\left|\frac{\mathrm{d}T}{\mathrm{d}t}\right|$, $|U|$, cl.cover, $\left|\frac{\mathrm{d}q}{\mathrm{d}t}\right|$

examples of parameter estimations





Important steps

- Generate a data set of model error and model tendencies to estimate the parameters γ, λ and σ .

  How to estimate the model error (tendencies)?

- Estimate the parameters of the error model.

  Estimation of diffusion coefficient λ seems to be most challenging.

- Solve the equation for model error tendency $\eta(t)$

  Fast and stable solution!!

Approach I (EM-Scheme)

$$\frac{\partial \eta_X(x,t)}{\partial t} = -\gamma(\tau_X)\eta_X(x,t) + \gamma(\tau_X)\nabla\cdot\left(\lambda^2(\tau_X)\nabla\eta_X(x,t)\right) + \sigma(\tau_X)\xi(x,t)$$

$\tau_X = \tau_X(x,t)$ is the tendency of the predictor variable X

∇^2 is a local operator $\rightarrow$ numerically efficient

García-Ojalvo et al., Generation of spatiotemporal noise, Phys. Rev. A 1992

Approach II (SPG – Stochastic pattern generator)

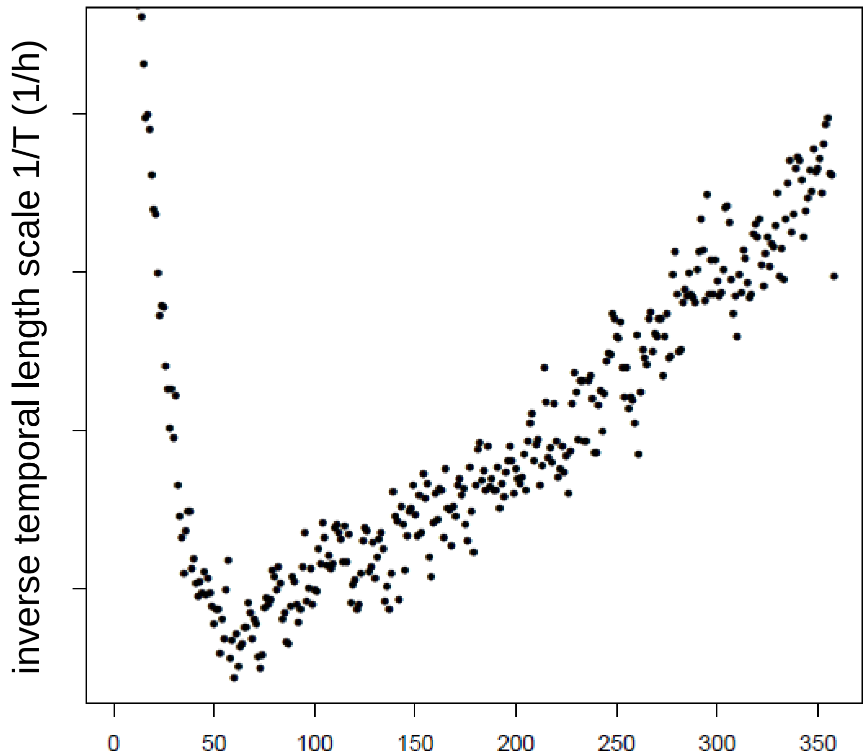
$$\left(\frac{\partial}{\partial t} + \gamma(\tau)\sqrt{1-\lambda^2(\tau)\nabla^2}\right)^3 \eta(x,t) = \sigma(\tau)\xi(x,t)$$

- fractional Laplacian for improved spatial characteristics
- finite variance of $\eta(t)$
- proportionality of scales
- non-local operator is numerically expensive for space-dependent coefficients γ, λ and σ

Tsyrulnikov & Gayfulin, A limited-area spatio-temporal stochastic pattern generator for simulation of uncertainties in ensemble applications, Met. Z. 2017

Proportionality of scales

modell error tendency field η for U (100m)



temporal length scale T_k
associated with
spatial wavenumber (or vector) k

$T_k \sim 1/k$ for $k \to \infty$

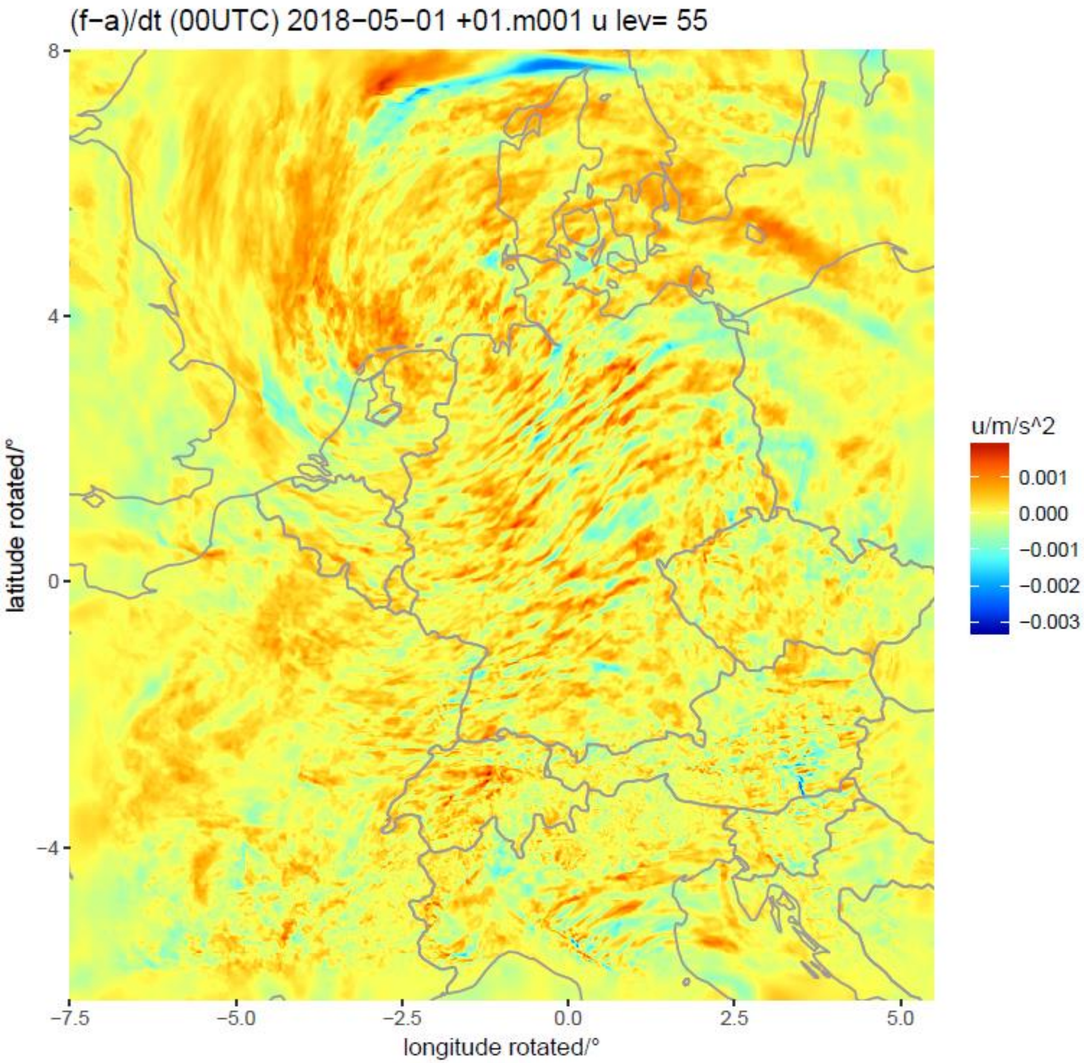
Comparison of the two approaches

	EM	SPG
Variance of η in continuous equation	infinite	finite
Proportionality of scales (strict)	no	yes
Diffusion operator	∇^2 local operator	Fractional Laplacian non-local operator
Computation on grid (physical space)	fast	(currently) expensive (ongoing task to find sufficiently accurate and affordable method)
Computation in Fourier Space	------	Fast for constant parameters γ, λ and σ

Parameter estimation for COSMO-D2-EPS

- Use 1-hour forecasts and COSMO-D2 analysis to estimate model error tendency

Model error tendency field based on 1h-forecast and analysis for U(300m)



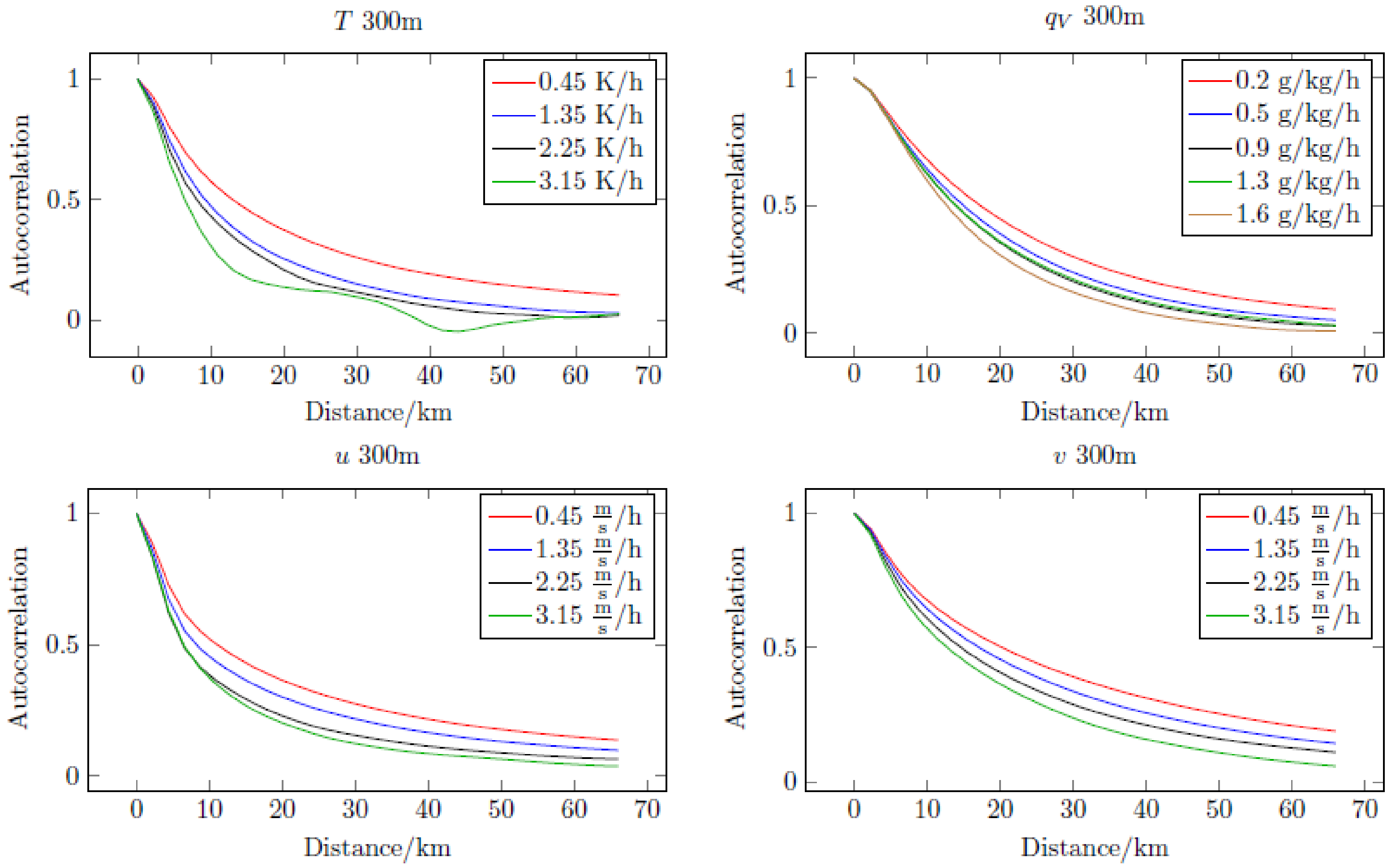
Parameter estimation for COSMO-D2-EPS

- Use 1-hour forecasts and COSMO-D2 analysis to estimate model error tendency

- Binning of data according to forecast tendency τ and for each bin
  - fit theoretical function to empirical autocorrelation $\rightarrow \lambda$
  - given λ, determine γ and σ from linear regression

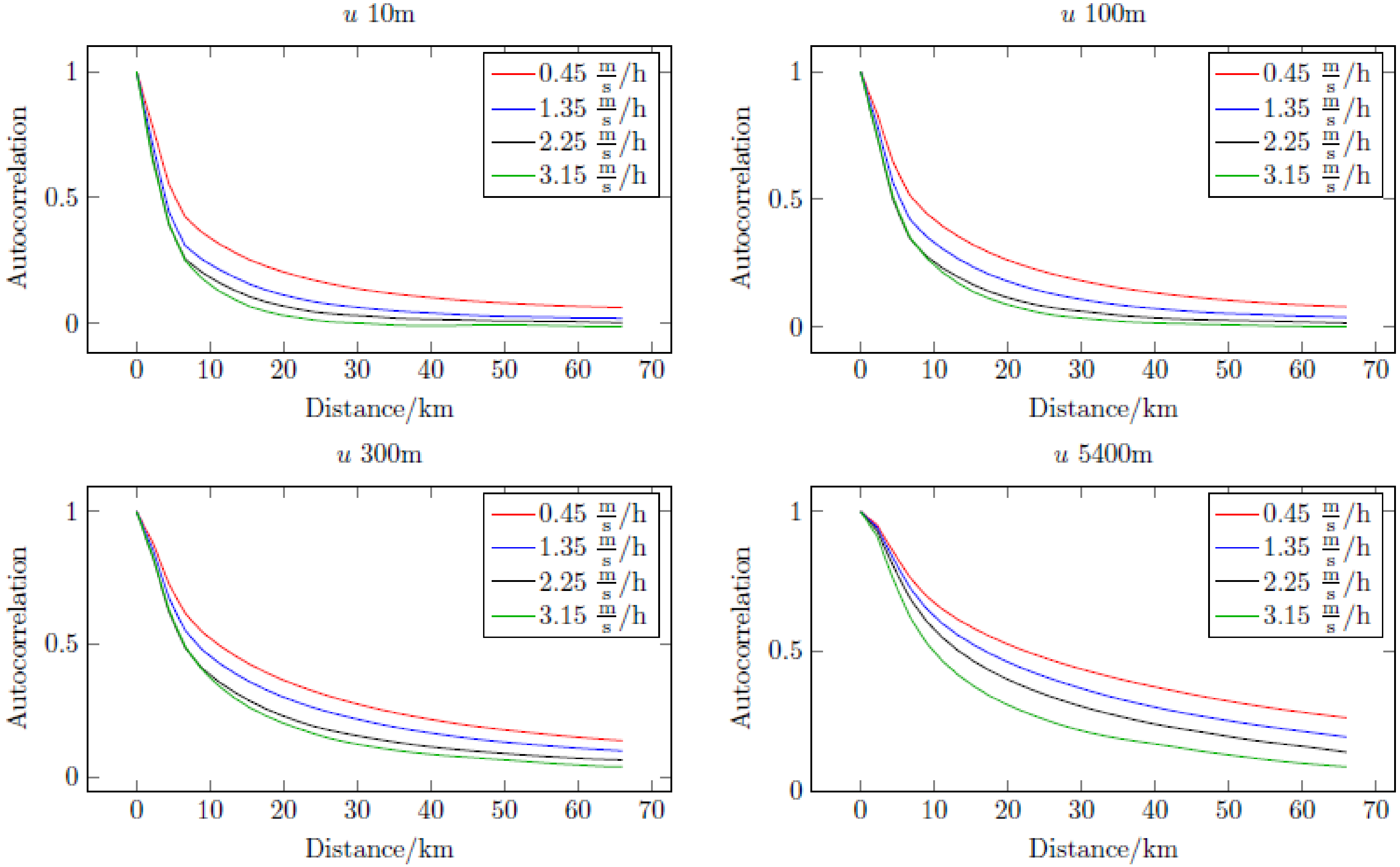
empirical autocorrelation in COSMO-D2

for different variables



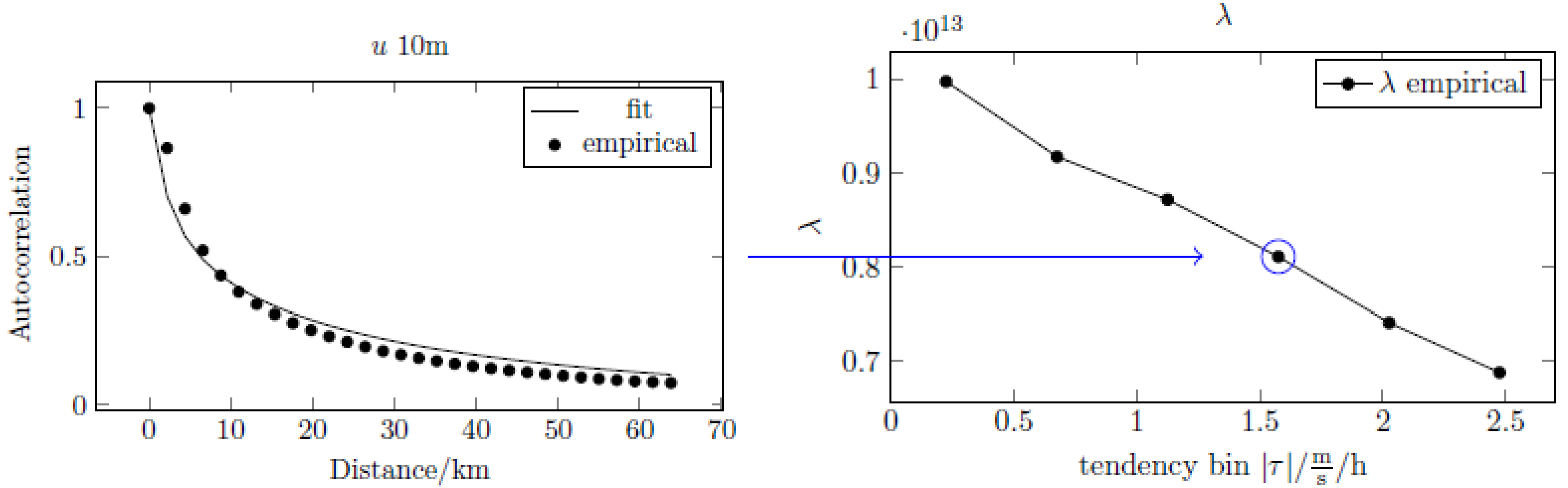
empirical autocorrelation in COSMO-D2

for different heights



Fit of empirical autocorrelation to functional form

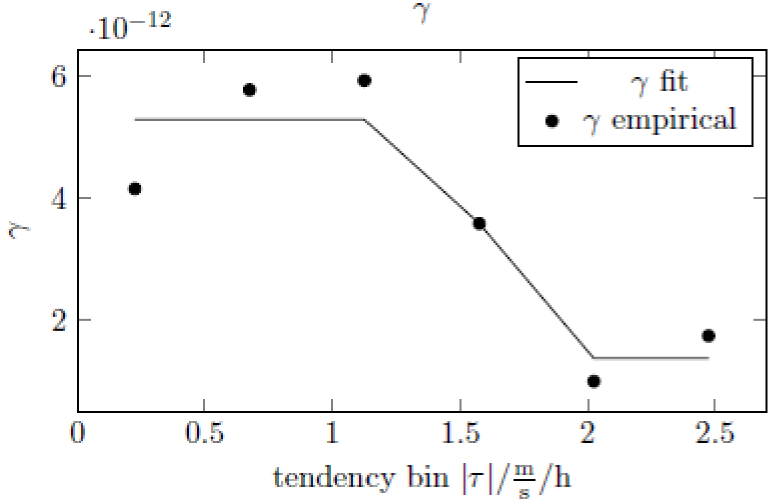
For each bin determine λ through fit:

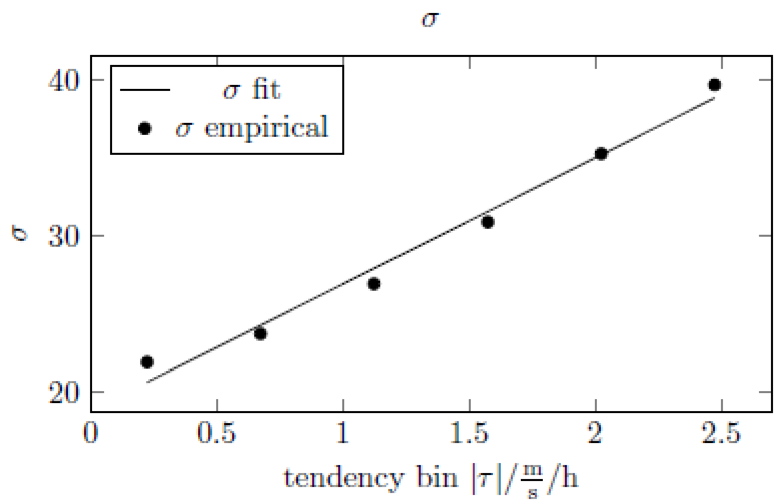


Assume some functional form for $\lambda(\tau)$ and fit to this function

Estimation of γ and σ by regression

Then determine γ and σ for each bin and assume some functional form





$$\gamma(\tau) = a_\gamma e^{-\left|\frac{\tau}{b_\gamma}\right|^{c_\gamma}} + d_\gamma$$
$$= \left(3.9 e^{-\left|4.4\cdot 10^4 \frac{\mathrm{s}^2}{\mathrm{m}}\tau\right|^{51}} + 1.4\right)\frac{10^{-12}}{\mathrm{s}}$$

$$\sigma(\tau) = a_\sigma|\tau| + b_\sigma$$
$$= 18.8\frac{|\tau|}{s} + 29181\frac{\mathrm{m}}{s^3}$$

Parameter estimation for COSMO-D2-EPS

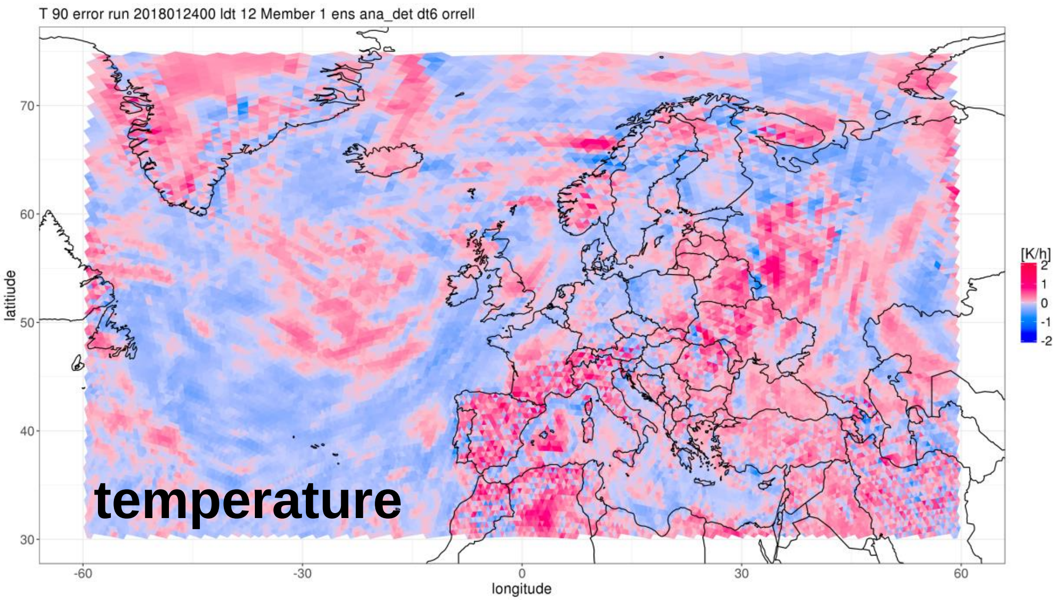
- Use 1-hour forecasts and COSMO-D2 analysis to estimate model error tendency

- Binning of data according to forecast tendency τ and for each bin
  - fit theoretical function to empirical autocorrelation $\rightarrow \lambda$
  - given λ, determine γ and σ from linear regression

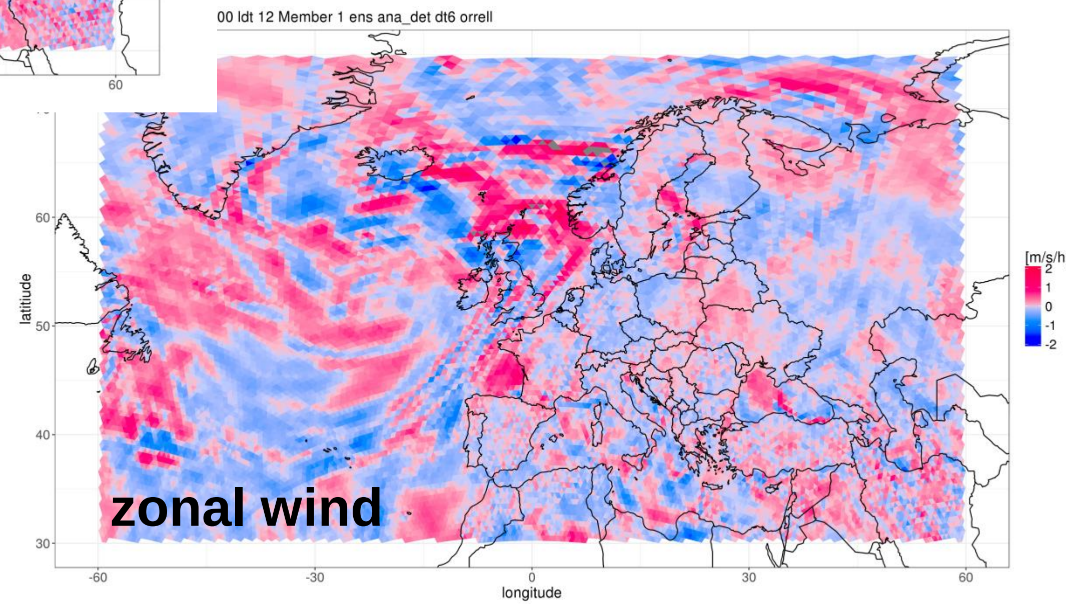
- However: estimation of λ is not stable against noise
  - Alternative: optimization with regularized cost function
  - this is computationally very expensive
  - result of first method can be used as initial guess for optimization

Parameter estimation for ICON-EPS

- Different options to estimate model error tendency
  - deterministic analysis as reference
    (13 km, coarse-grained to 40 km)
  - EDA mean as reference (40km as for ICON-EPS)
  - older ICON-EPS forecast run as reference (40km)



Model error tendency field based on 6h forecast / analysis differences on lowest model level

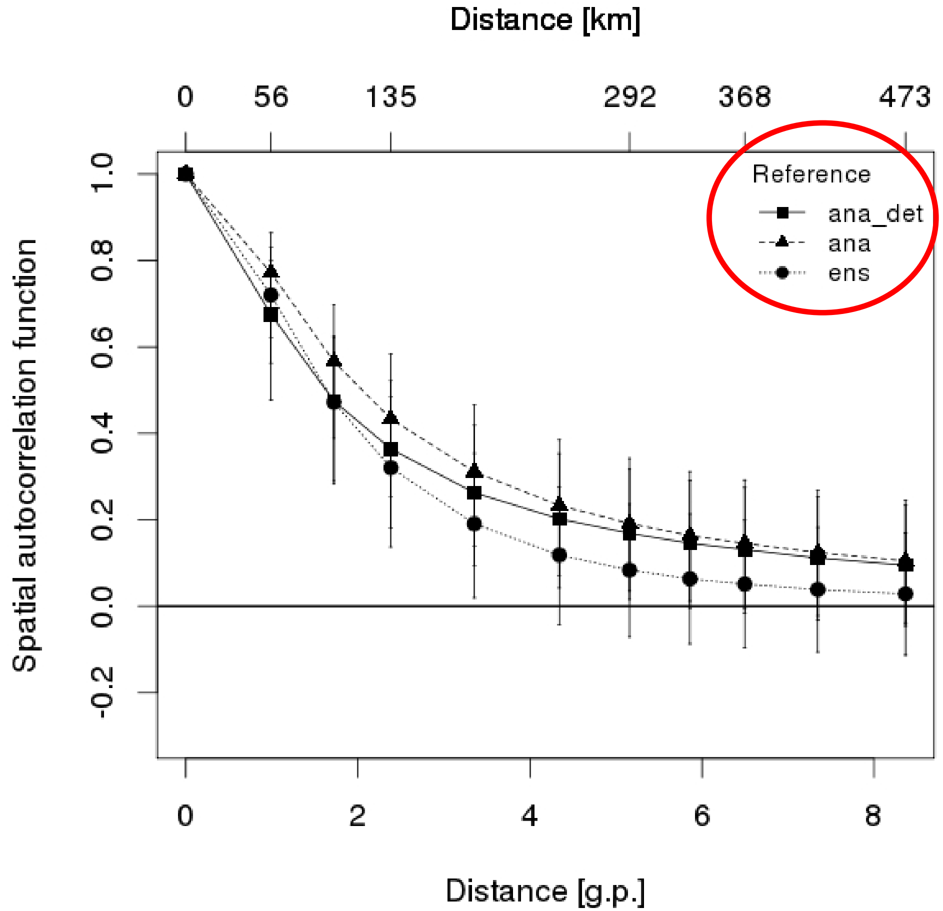


Parameter estimation for ICON-EPS

- Different options to estimate model error tendency
  - deterministic analysis as reference
    (13 km, coarse-grained to 40 km)
  - EDA mean as reference (40km as for ICON-EPS)
  - older ICON-EPS forecast run as reference (40/20km)

- Binning of data according to forecast tendency τ and for each bin
  - fit theoretical function to empirical autocorrelation $\rightarrow \lambda$
  - given λ, determine γ and σ from linear regression

empirical autocorrelation in ICON
for different error reference fields



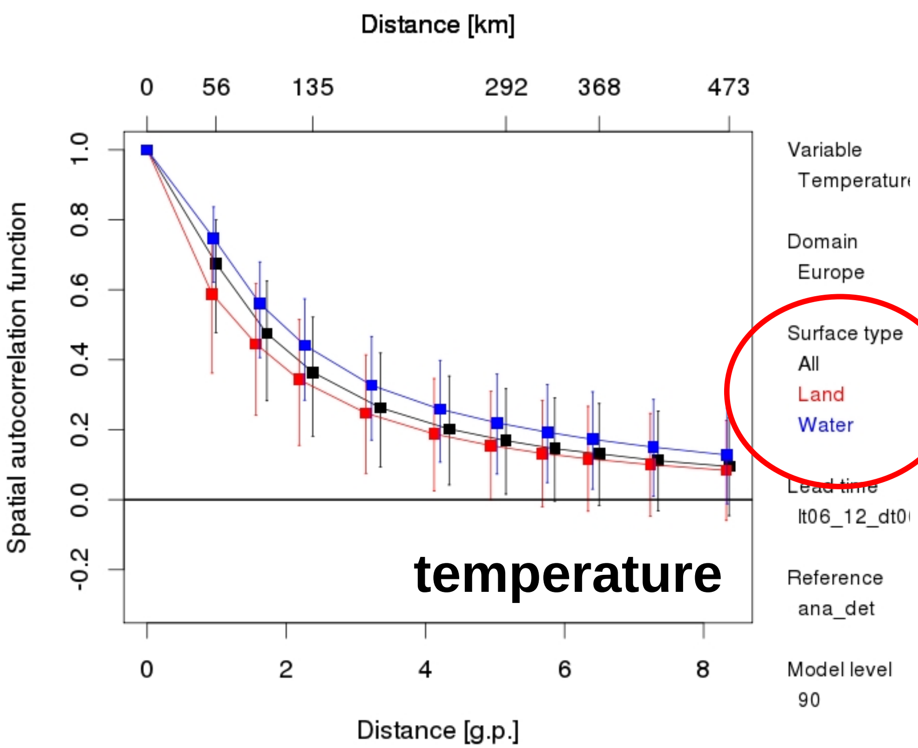
temperature
on lowest model level
Europe

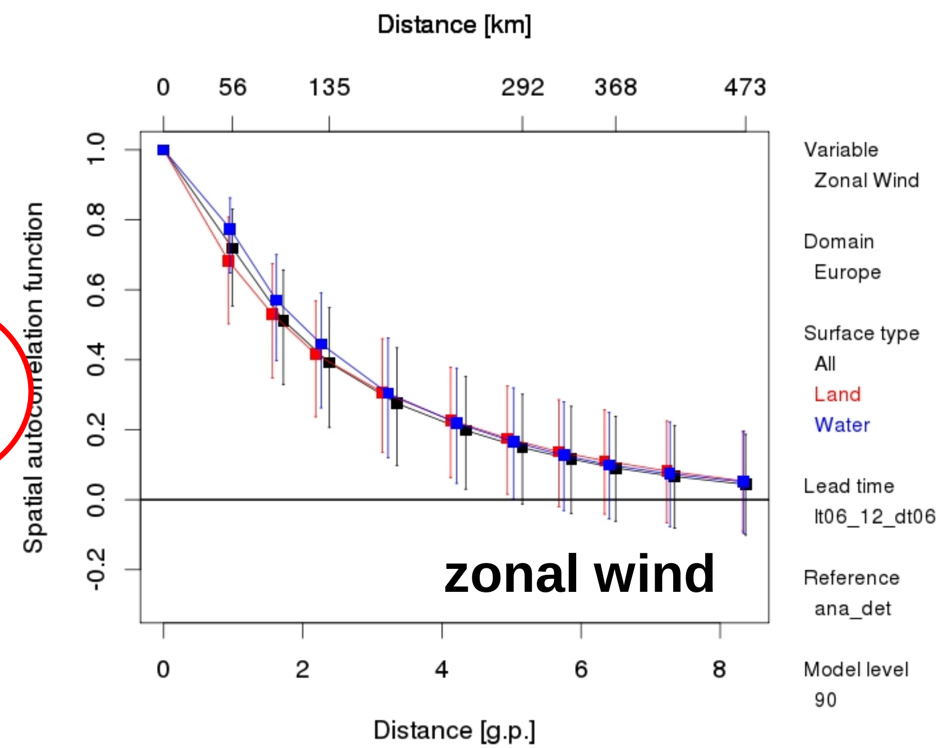
Lead time
lt06_12_dt06

Model level
90

empirical autocorrelation in ICON

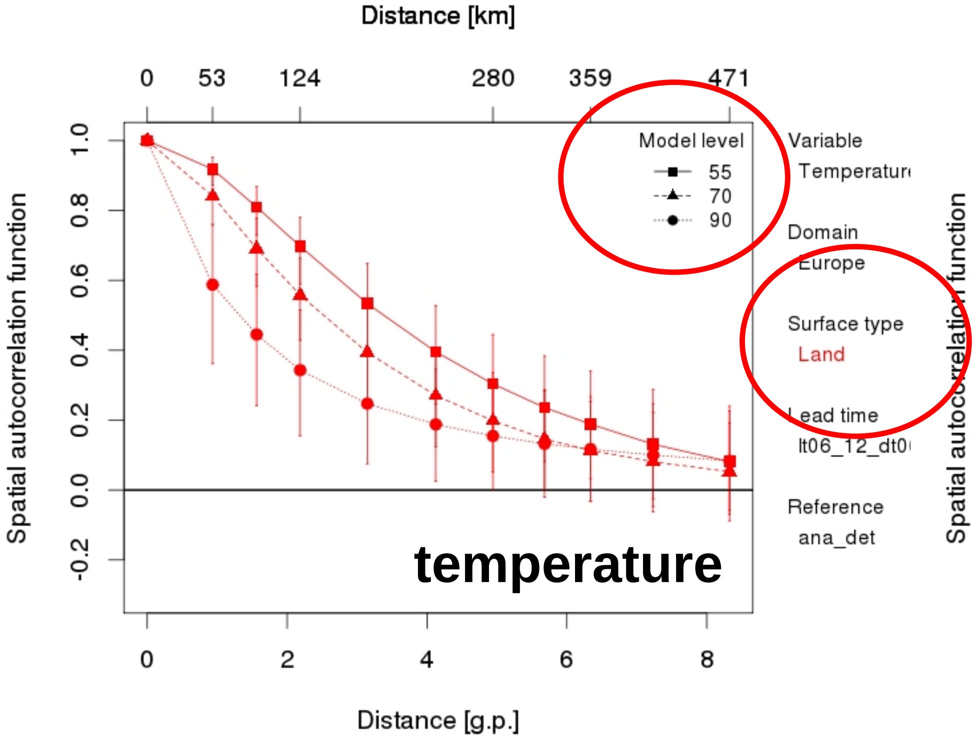
for different variables & surface types

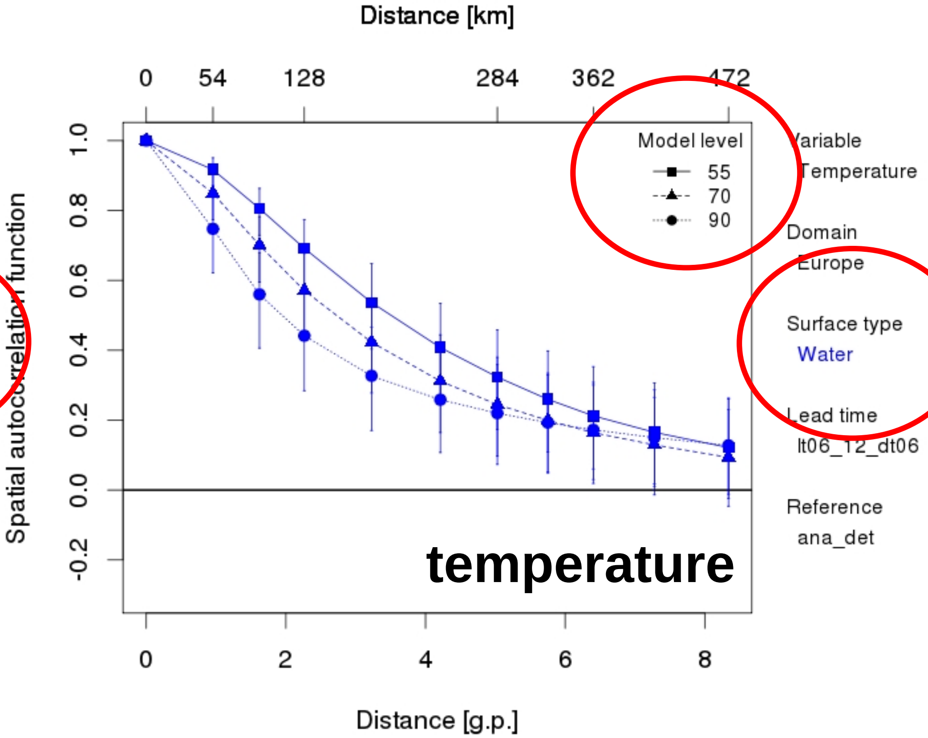




empirical autocorrelation in ICON

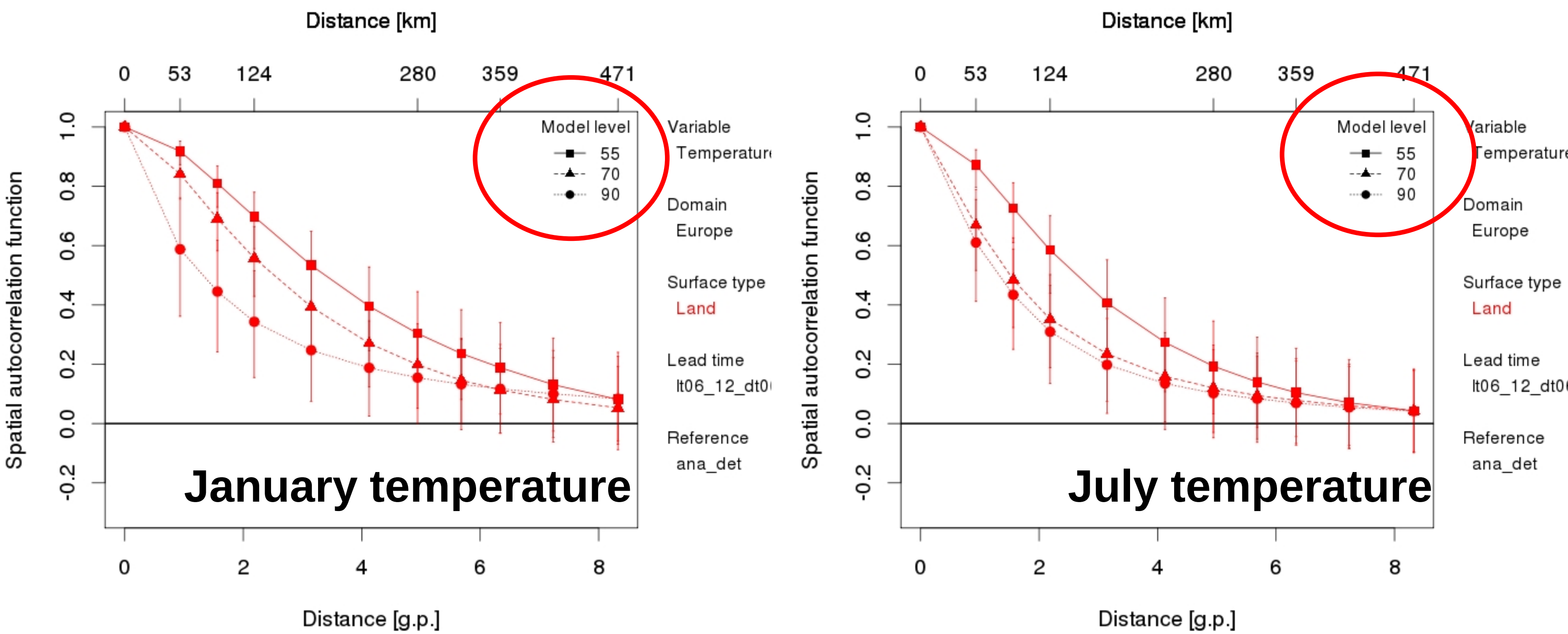
for different model levels & surface types





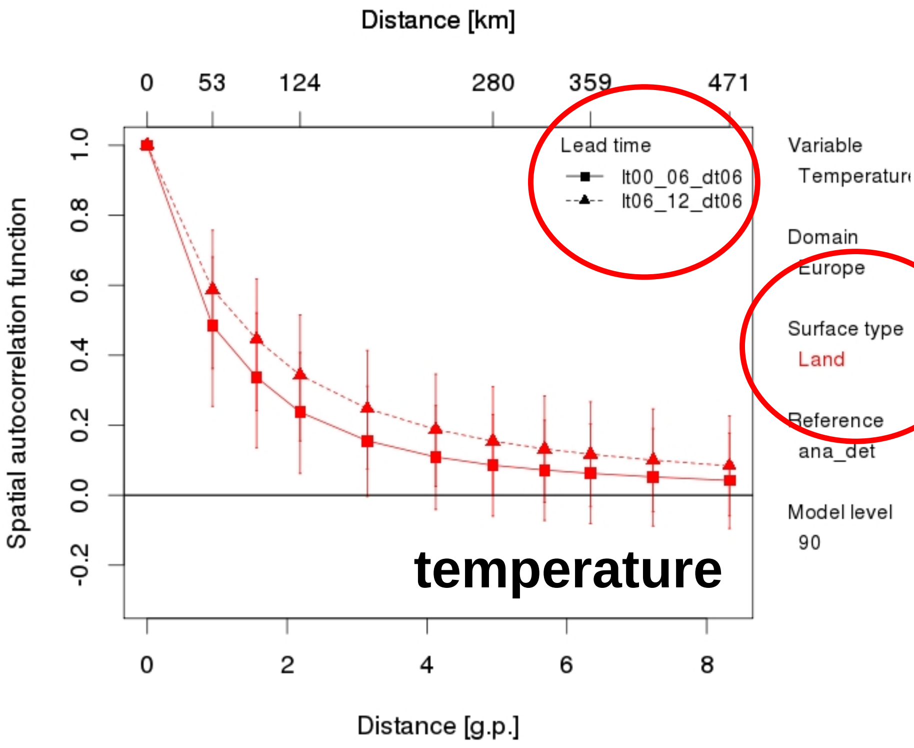
empirical autocorrelation in ICON

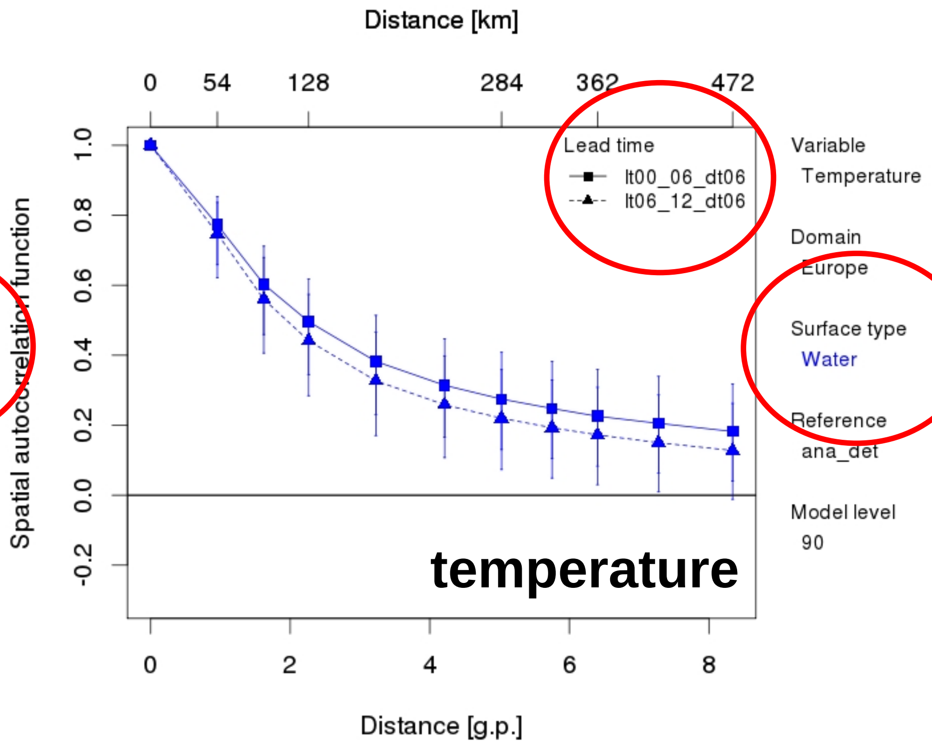
for different seasons & model levels



empirical autocorrelation in ICON

for different lead times of error tendency estimation & surface types





empirical autocorrelation of model error tendency in ICON – summary

- Dependent on surface type, model height and variable (with interactions), less dependent on the season and the forecast lead time.
- Near the surface, model error tendency approximations of temperature have longer correlations lengths with analyses as reference than with forecasts as reference.
- With increasing model height, model error tendency approximation is less dependent on the reference dataset for error estimation.
- Higher dependence on surface type for temperature than for zonal wind.

Open questions and next steps

COSMO-D2

- further research on SPG (fast computation for space-dependent parameters, use regionally constant parameters…)
- decide on approach for estimation of model parameters
- run first long period experiments

ICON

- further research on appropriate reference fields for the estimation of model error tendencies
- implementation of stochastic differential equation for the model error in ICON
- estimation of model parameters
- run experiments

Thank you for your attention!!

CONSORTIUM FOR SMALL SCALE MODELING
COSMO