

# From Catchments as Organised Systems to Models Based on Functional Units

Final Report and Synthesis, December 2021

## DFG PI from Germany

 Karlsruhe Institute of Technology (KIT), Institute for Water and River Basin Management, Chair for Hydrology (Prof. E. Zehe, Dr. U. Ehret), Institute for Photogrammetry and Remote Sensing, (Prof. S. Hinz)	Karlsruhe Institute of Technology (KIT), Institute for Water and River Basin Management, Chair for Hydrology (Prof. E. Zehe, Dr. U. Ehret), Institute for Photogrammetry and Remote Sensing, (Prof. S. Hinz)
	University of Freiburg, Institute of Hydrology (Prof. M. Weiler)
	Helmholtz-Centre Potsdam, GFZ German Research Centre for Geosciences, Section Hydrology (Dr. T. Blume)
	Max Planck Institute for Biogeochemistry Jena, Biospheric Theory and Modelling Group (Dr. A. Kleidon)
	Universität Hohenheim, Institute of Physics and Meteorology (Prof. V. Wulfmeyer)
	University of Potsdam, Institute of Earth- and Environmental Sciences (Prof. J. Tronicke)
	Technische Universität Braunschweig, Institute of Geoecology (Prof. B. Schröder-Esselbach, Dr. Loes van Schaik)

## FNR/DFG Lead Agency PI from Luxembourg

	Luxembourg Institute of Science and Technology. (Prof. Dr. habil. L. Pfister, Dr. C. Hissler, Dr. M. Schlerf, Dr. K. Mallik, Dr. P. Matgen, Dr. I. Trebs)
---	---

## FWF/DFG Lead Agency PI from Austria

 Universität für Bodenkultur Wien	University of Natural Resources and Life Sciences (BOKU) Vienna, Institute of Water Management, Hydrology and Hydraulic Engineering (IWHW) (Univ.Prof. Karsten Schulz)
---	--

1	General Information.....	5
1.1	DFG reference number .....	5
1.2	Scientific speakers .....	5
1.3	Report- and funding period .....	5
1.4	List of the most relevant publications.....	5
1.4.1	CAOS Special Issues .....	5
1.4.2	Peer-reviewed journal articles .....	6
1.4.2.1	Phase II Project A and phase I project J: PI B. Schröder, J. Eccart., L.van Schaik, Team: K. Schneider, A. Zangerlé, T. Hohenbrink .....	6
1.4.2.2	Phase II Project B 1 (geophysics) and phase I project E: PI J. Tronicke, Team: N. Allroggen .....	6
1.4.2.3	Phase II Project B 2 (remote sensing), PI S. Hinz, Team: S. Keller.....	7
1.4.2.4	Phase II Project C1 and phase II project B. PI Karsten Schulz, Team. B Müller, C. Brenner, E Thiem. M. Schlerf. K. Mallick.....	7
1.4.2.5	Phase I project C2 and phase I project C: PI Wulfmeyer, Team: S. Bauer,.....	7
1.4.2.6	Phase II Project C3 and C5 and phase I Project D, PI A. Kleidon, Team: M. Renner .....	8
1.4.2.7	Phase II project D and phase I project A, PI L. Pfister, Team F. Fenicia, N. Martínez- Carreras, C. Hissler, L. Godol .....	8
1.4.2.8	Phase II Project D and phase I project H, PI M. Weiler, T.Blume. Team. M Sprenger, L. Angermann, D. Demand.....	9
1.4.2.9	Phase II Project E and phase I Project S, PI U. Ehret, Team: M. Westhoff, M. Neuper, R. van Pruissen .....	9
1.4.2.10	Phase II Project F and phase I Project I, PI E. Zehe, CoPi Th. Blume Team: R. Loritz, C. Jackisch.....	10
1.4.2.11	Phase I and phase II Project G, PI T. Blume, M. Weiler, Team: S. Hassler, D. Beiter, E. Sohr (geb. Lieder), N. Kaplan .....	11
1.4.3	Most important CAOS Conferences in the funding period .....	11
2	Final progress report .....	12
2.1	CAOS research questions, hypotheses and structure .....	12
2.1.1	First funding period .....	12
2.1.2	Second funding period.....	12
2.2	Project development .....	16
2.2.1	Experimental characterisation of landscape structure and functioning .....	16

2.2.1.1	The CAOS monitoring network to characterise landscape structure and functional units	16
2.2.1.1.1	Conceptual framework and selection of candidate functional units .....	16
2.2.1.1.2	Selection and instrumentation of candidate elementary functional units .....	18
2.2.1.1.3	Synoptic assessment of landscape structure and similarity .....	20
2.2.1.2	Characterising sub structure and preferential flow across scales and geological settings	22
2.2.1.2.1	Combined tracer tests, soil survey and earthworm sampling .....	22
2.2.1.2.2	In situ geophysical imaging of subsurface structures and flow processes .....	24
2.2.1.2.3	Identification of preferential flow occurrence through soil moisture monitoring at 135 soil profiles .....	29
2.2.1.3	Characterizing water storage, mixing and stream flow release.....	30
2.2.1.3.1	Soil moisture dynamics across scales and landscape settings .....	30
2.2.1.3.2	Travel/ transit time distribution of runoff components across scales .....	33
2.2.1.3.3	Riparian exchange and hydrological connectivity.....	38
2.2.1.3.4	Seasonality and storage controls of stream flow generation .....	44
2.2.1.4	Radar based quantitative rainfall estimates .....	46
2.2.1.5	Characterizing land surface atmosphere energy exchange and feedbacks .....	48
2.2.1.5.1	Monitoring and experimental campaigns across scales .....	48
2.2.1.5.2	Tree transpiration inferred from sap flow velocities: variability and uncertainty ...	49
2.2.1.5.1	Inferring root water uptake from soil moisture observations .....	51
2.2.1.5.2	Detecting sensible and latent heat fluxes using unmanned aerial systems .....	53
2.2.1.5.3	Diurnal variation and hysteresis patterns of evapo-transpiration as means of model discrimination .....	53
2.2.1.6	Synthesizing minimum data needs for the identification of functional units .....	56
2.2.2	Key model and theory developments .....	61
2.2.2.1	Functional units for runoff generation and dynamic hydrological similarity .....	61
2.2.2.1.1	Modelling lower mesoscale catchments with representative hillslopes .....	61
2.2.2.1.2	The dynamic nature of functional similarity.....	63
2.2.2.2	Adaptive models for the water and energy cycle.....	65
2.2.2.2.1	The CAOS model – an adaptive model for the water and energy balance of lower mesoscale catchments.....	65

2.2.2.2.2	Adaptive model clustering: improving modelling efficiency by using dynamic functional similarity .....	68
2.2.2.3	New model approaches for preferential flow .....	72
2.2.2.3.1	A Lagrangian framework for soil water dynamics and preferential solute transport 72	
2.2.2.3.2	Testing the Stokes flow approach for simulation infiltration across scales and landscape settings .....	75
2.2.2.4	Improving quantitative rainfall estimates and simulations of land surface atmosphere energy exchange fluxes with WRF .....	76
2.2.2.4.1	Modell based quantitative rainfall estimates by assimilation of weather radar .....	76
2.2.2.4.2	Coupled simulations of land surface atmosphere energy exchange fluxes with WRF-NOAH-MP-HYDRO.....	78
2.2.2.5	A thermodynamic foundation of the hydrological cycle and hydrological modelling .	80
2.2.2.5.1	Explaining and predicting the surface energy balance, evaporation and atmospheric water cycling in a thermodynamic framework .....	80
2.2.2.5.2	Explaining and modelling structural controls on rainfall runoff, preferential flow and storage dynamics in a thermodynamic framework .....	84
2.2.2.6	Information as universal measure to quantify uncertainty, similarity and organisation 90	
2.3	Central issues and research data management .....	94
2.4	Qualification of young researchers in CAOS .....	94
3	Summary and conclusions .....	96
3.1	Q1: How to define hydrological functional similarity and its relation to landscape structure? D1: Connect landscape organisation and hydrological functioning.....	96
3.2	Q2: Does spatial organisation in catchments imply the existence of functional units? Functional units/ D2: Experimental protocol to characterise structure and functioning. ....	98
3.3	Q3: How to balance necessary model complexity and falsifiability? D3: Advancements in model concepts and new framework for model inter comparison using information theory.....	100
3.4	Q4: What is the connection between thermodynamics, organising principles, catchment structure and functioning? D4: Protocol for testing organizing principles.....	102
4	Publication of data from final reports.....	104
5	References not included in section 1.4 .....	104

# 1 General Information

## 1.1 DFG reference number

Research Unit FOR 1598

## 1.2 Scientific speakers

<p>Prof. Dr. Erwin Zehe (Phase I and II) Karlsruhe Institute of Technology Institute of Water and River Basin Management Kaiserstr. 12, 76131 Karlsruhe 76131 Karlsruhe Deutschland phone: +49 721 608-43814 e-mail: <a href="mailto:Erwin.Zehe@kit.edu">Erwin.Zehe@kit.edu</a></p>	<p>Prof. Dr. habil. L. Pfister (Phase II) 'Catchment and eco-hydrology' research group Department 'Environmental Research and Innovation', Luxembourg Institute of Science and Technology 41, rue du Brill L-4422 Belvaux Grand-Duchy of Luxembourg phone: +27 58 88 460 e-mail: <a href="mailto:Laurent.Pfister@list.lu">Laurent.Pfister@list.lu</a></p>
<p>Dr. Theresa Blume (Phase I) Section Hydrology, German Research Centre for Geosciences Potsdam Telegrafenberg, Building C 4, 14473 Potsdam. phone: +49 331 288-1512 email: <a href="mailto:theresa.blume@gfz-potsdam.de">theresa.blume@gfz-potsdam.de</a></p>	<p>Prof. Dr. Markus Weiler (Phase I) Institute of Hydrology, University of Freiburg, Friedrichstraße 39 79098 Freiburg phone : +49 761 - 203 3535 email: <a href="mailto:markus.weiler@hydrology.uni-freiburg.de">markus.weiler@hydrology.uni-freiburg.de</a></p>

## 1.3 Report- and funding period

01.01.2012 – 31.12.2014 (phase I), 01.01.2015 - 31.12.2019 (phase II)

## 1.4 List of the most relevant publications

### 1.4.1 CAOS Special Issues

- *Linking landscape organisation and hydrological functioning: from hypotheses and observations to concepts, models and understanding.* Joint SI Hydrology and Earth System Sciences and Earth System Science Data. Guest Editors: L. van Schaik, T. Hohenbrink, C. Jackisch, H. Laudon, L. Pfister, S.K. Hassler, M. Renner, H. McMillan, T. Blume, P. Gentine, and P. Saco.  
[https://www.hydrol-earth-syst-sci.net/special\\_issue985.html](https://www.hydrol-earth-syst-sci.net/special_issue985.html)
- *Thermodynamics and optimality in the Earth system and its subsystems.* Joint SI Earth System Dynamics and Hydrology and Earth System Sciences. Guest Editors: A. Kleidon, P. Cox, H. Savenije, E. Zehe, M. Crucifix, and S. Hergarten.  
[https://www.hydrol-earth-syst-sci.net/special\\_issue430\\_934.html](https://www.hydrol-earth-syst-sci.net/special_issue430_934.html)
- *Towards theories that link catchment structures and model structures.* SI Hydrology and Earth System Sciences. Guest Editors: E. Zehe, S. Schymanski, G. Blöschl, H. Gupta, M. Sivapalan. [https://www.hydrol-earth-syst-sci.net/special\\_issue132.html](https://www.hydrol-earth-syst-sci.net/special_issue132.html)

## 1.4.2 Peer-reviewed journal articles

### 1.4.2.1 Phase II Project A and phase I project J: PI B. Schröder, J. Eccart., L.van Schaik, Team: K. Schneider, A. Zangerlé, T. Hohenbrink

Dormann, C.F., Guillera-Aroita, G., Calabrese, J., Matechou, E., Barton, K., Bahn, V., Beale, C.M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J.J., Lennon, J., Pollock, L.S., Reineking, B., Roberts, D., **Schröder**, B., Thuiller, W., Warton, D.I., Wintle, B.A., Wood, S., Wüest, R., Hartig, F. 2018. Model averaging in ecology: a review of Bayesian, information-theoretic, machine-learning and other approaches. *Ecological Monographs*, 88(4): 485-504.

Phillips, H. R. P., Guerra, C. A., Bartz, M. L. C., Briones, J. I., Brown, G., Crowther, T. W., Ferlian, O., Gongalsky, K. B., van den Hoogen, J., Krebs, J., Orgiazzi, A., Routh, D., Schwarz, B., Bach, E. M., Bennett, J., Brose, U., Decaens, T., König-Ries, B., Loreau, M., Mathieu, J., Mulder, C., van der Putten, W. H., Ramirez, K. S., Rillig, M. C., Russell, D., Rutgers, M., Thakur, M. P., de Vries, F. T., Wall, D. H., Wardle, D. A., Arai, M., Ayuke, F. O., Baker, G. H., Beausejour, R., Bedano, J. C., Birkhofer, K., Blanchart, E., Blossey, B., Bolger, T., Bradley, R. L., Callahan, M. A., Capowiez, Y., Caulfield, M. E., Choi, A., Crotty, F. V., Davalos, A., Cosin, D. J. D., Dominguez, A., Duhour, A. E., van Eekeren, N., Emmerling, C., Falco, L. B., Fernandez, R., Fonte, S. J., Fragoso, C., Franco, A. L. C., Fugere, M., Fusilero, A. T., Gholami, S., Gundale, M. J., Lopez, M. G., Hackenberger, D. K., Hernandez, L. M., Hishi, T., Holdsworth, A. R., Holmstrup, M., Hopfensperger, K. N., Lwanga, E. H., Huhta, V., Hurisso, T. T., Iannone, B. V., Iordache, M., Joschko, M., Kaneko, N., Kanianska, R., Keith, A. M., Kelly, C. A., Kernecker, M. L., Klaminder, J., Kone, A. W., Kooch, Y., Kukkonen, S. T., Lalthanzara, H., Lammel, D. R., Lebedev, I. M., Li, Y., Lidon, J. B. J., Lincoln, N. K., Loss, S. R., Marichal, R., Matula, R., Moos, J. H., Moreno, G., Moron-Rios, A., Muys, B., Neirynek, J., Norgrove, L., Novo, M., Nuutinen, V., Nuzzo, V., Rahman, P. M., Pansu, J., Paudel, S., Peres, G., Perez-Camacho, L., Pineiro, R., Ponge, J. F., Rashid, M. I., Rebollo, S., Rodeiro-Iglesias, J., Rodriguez, M. A., Roth, A. M., Rousseau, G. X., Rozen, A., Sayad, E., **van Schaik, L.**, Scharenbroch, B. C., Schirrmann, M., Schmidt, O., **Schröder, B.**, Seeber, J., Shashkov, M. P., Singh, J., Smith, S. M., Steinwandter, M., Talavera, J. A., Trigo, D., Tsukamoto, J., de Valenca, A. W., Vanek, S. J., Virto, I., Wackett, A. A., Warren, M. W., Wehr, N. H., Whalen, J. K., Wironen, M. B., Wolters, V., Zenkova, I. V., Zhang, W., Cameron, E. K., and Eisenhauer, N.: Global distribution of earthworm diversity (vol 366, pg 480, 2019), *Science*, 369, 3, 10.1126/science.abd9834, 2020.

Palm J, **van Schaik NLMB, Schröder B.**: Modelling distribution patterns of anecic, epigeic and endogeic earthworms at catchment-scale in agro-ecosystems. *Pedobiologia* 56(1): 23-31, 2013.

**Reck, A., Jackisch, C., Hohenbrink, T. L., Schröder, B., Zangerlé, A., van Schaik, L.**: Impact of Temporal Macropore Dynamics on Infiltration: Field Experiments and Model Simulations. *Vadose Zone J*, 17:170147, 2018.

**Schneider, A.-K., van Schaik, L., Zangerlé, A., Schröder, B.**: Which abiotic filters shape earthworm distribution patterns at catchment scale? *Eur J Soil Sci*, 67: 431-442, 2016.

**Schneider A.-K., Schröder B** : Perspectives in modelling earthworm dynamics and their feedbacks with abiotic soil properties, *Applied Soil Ecology*, 58, 29-36, 2012.

**Schneider, A. K., Hohenbrink, T. L., Reck, A., Zangerle, A., Schröder, B., Zehe, E., and van Schaik, L.**: Variability of earthworm-induced biopores and their hydrological effectiveness in space and time, *Pedobiologia*, 71, 8-19, 10.1016/j.pedobi.2018.09.001, 2018.

**Van Schaik NLMB, Palm J, Klaus J, Zehe E, Schröder B.**: Linking spatial earthworm distribution to macropore numbers and hydrological effectiveness, *Ecohydrology*, 7: 401-408, 2014.

**Van Schaik, N.L.M.B., Palm, J., Klaus, J., Zehe, E., Schröder, B.**: Potential effects of tillage and field borders on within field spatial distribution patterns of earthworms. *Agr Ecosyst Environ*, 228: 82 -90, 2016.

**Zangerlé, A., Hissler, C., van Schaik, L., McKey, D.**: Identification of earthworm burrow origins by near infrared spectroscopy: combining results from field sites and laboratory microcosms. *Soil Till Res*, 155: 280-288, 2016.

### 1.4.2.2 Phase II Project B 1 (geophysics) and phase I project E: PI J. Tronicke, Team: N. Allroggen

**Allroggen N., Tronicke J., Delock M. and Boniger U.**: Topographic migration of GPR data with variable velocities. In IEEE conference publications: Advanced Ground Penetrating Radar (IWAGPR), 2013 7th International Workshop on Ground Penetrating Radar. 59–63, doi:10.1109/IWAGPR.2013.6601514, 2013.

**Allroggen, N., van Schaik, N.L.M.B., Tronicke, J.:** 4D ground-penetrating radar during a plot scale dye tracer experiment. *Journal of Applied Geophysics*, vol. 118, pp. 139–144, 2015.

**Allroggen, N., Tronicke, J., Delock, M., and Boniger, U.:** Topographic migration of 2D and 3D ground-penetrating radar data considering variable velocities, *Near Surface Geophysics*, 13, 10.3997/1873-0604.2014037, 2015.

**Allroggen, N., and Tronicke, J.:** Attribute-based analysis of time-lapse ground-penetrating radar data, *Geophysics*, 81, H1-H8, 10.1190/geo2015-0171.1, 2016.

**Allroggen, N., Jackisch, C., Tronicke, J.:** Four-dimensional gridding of time-lapse GPR data. In *Proceedings of 9th International Workshop on Advanced Ground Penetrating Radar*. Edinburgh, Scotland, 2017.

**Allroggen, N., Beiter, D., and Tronicke, J.:** Ground-penetrating radar monitoring of fast subsurface processes, *Geophysics*, 85, A19-A23, 10.1190/geo2019-0737.1, 2020.

1.4.2.3 Phase II Project B 2 (remote sensing), PI S. Hinz, Team: S. Keller

**Keller, S., F. M. Riese, N. Allroggen, C. Jackisch, and S. Hinz,** Modeling Subsurface Soil Moisture Based on Hyperspectral Data: First Results of a Multilateral Field Campaign, vol. 37, pp. 34–48, *Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation*, München, 2018.

1.4.2.4 Phase II Project C1 and phase II project B. PI Karsten Schulz, Team. B Müller, C. Brenner, E Thiem. M. Schlerf. K. Mallick

**Brenner, Claire, Christina Elisabeth Thiem, Hans Dieter Wizemann, Matthias Bernhardt, and Karsten Schulz.:** Estimating Spatially Distributed Turbulent Heat Fluxes from High-Resolution Thermal Imagery Acquired with a UAV System. *International Journal of Remote Sensing* 38 (8–10): 3003–26. <https://doi.org/10.1080/01431161.2017.1280202>, 2017.

**Brenner, Claire, Matthias Zeeman, Matthias Bernhardt, and Karsten Schulz.:** Estimation of Evapotranspiration of Temperate Grassland Based on High-Resolution Thermal and Visible Range Imagery from Unmanned Aerial Systems. *International Journal of Remote Sensing* 39 (15–16): 5141–74. <https://doi.org/10.1080/01431161.2018.1471550>, 2018.

**Mallick, K., Trebs, I., Boegh, E., Giustarini, L., Schlerf, M., Drewry, D. T., Hoffmann, L., von Randow, C., Kruijt, B., Araùjo, A., Saleska, S., Ehleringer, J. R., Domingues, T. F., Ometto, J. P. H. B., Nobre, A. D., de Moraes, O. L. L., Hayek, M., Munger, J. W., and Wofsy, S. C.:** Canopy-scale biophysical controls of transpiration and evaporation in the Amazon Basin, *Hydrol. Earth Syst. Sci.*, 20, 4237–4264, <https://doi.org/10.5194/hess-20-4237-2016>, 2016.

**Müller, Benjamin, Matthias Bernhardt, Conrad Jackisch, and Karsten Schulz:** Estimating Spatially Distributed Soil Texture Using Time Series of Thermal Remote Sensing - A Case Study in Central Europe." *Hydrology and Earth System Sciences* 20 (9): 3765–75. <https://doi.org/10.5194/hess-20-3765-2016>, 2016.

**Müller, Benjamin, Matthias Bernhardt, and Karsten Schulz:** Identification of Catchment Functional Units by Time Series of Thermal Remote Sensing Images. *Hydrology and Earth System Sciences* 18 (12): 5345–59. <https://doi.org/10.5194/hess-18-5345-2014>, 2014.

**Sun, Liya, and Karsten Schulz:** Response to Johnson B.A. Scale Issues Related to the Accuracy Assessment of Land Use/Land Cover Maps Produced Using Multi-Resolution Data: Comments on 'The Improvement of Land Cover Classification by Thermal Remote Sensing'. *Remote Sens.* 2015, 7, 8368-839. *Remote Sensing* 7 (10): 13436–39. <https://doi.org/10.3390/rs71013436>, 2015a.

**Sun, Liya, and Karsten Schulz:** The Improvement of Land Cover Classification by Thermal Remote Sensing. *Remote Sensing* 7 (7): 8368–90. <https://doi.org/10.3390/rs70708368>, 2015b.

**Sun, Liya, and Karsten Schulz:** Spatio-Temporal LAI Modelling by Integrating Climate and MODIS LAI Data in a Mesoscale Catchment. *Remote Sensing* 9 (2): 1–20. <https://doi.org/10.3390/rs9020144>, 2017.

1.4.2.5 Phase I project C2 and phase I project C: PI Wulfmeyer, Team: S. Bauer,

**Bauer, H.-S., T. Schwitalla, V. Wulfmeyer, A. Bakhshaii, U. Ehret, M. Neuper, and O. Caumont,** Quantitative precipitation estimation based on high-resolution numerical weather prediction and data assimilation with WRF – a performance test, *Tellus A Dyn. Meteorol. Oceanogr.*, 67(1), 25047. doi:10.3402/tellusa.v67.25047; 2015.

Greve, P., K. Warrach-Sagi, and V. **Wulfmeyer**: Evaluating soil water content in a WRF-NOAH downscaling experiment. *J. Appl. Meteorol. Climatol.* 52, 2312–2327, DOI: 10.1175/JAMC-D-12-0239.1, 2013.

Schwitalla, T., and V. **Wulfmeyer**: Radar data assimilation experiments using the IPM WRF Rapid Update Cycle. *Meteorol. Z.*, DOI:10.1127/0941-2948/2014/0513, 2014.

#### 1.4.2.6 Phase II Project C3 and C5 and phase I Project D, PI A. Kleidon, Team: M. Renner

Conte, L., **Renner**, M., Brando, P., Santos, C. O., Silverio, D., Kolle, O., Trumbore, S. E. and **Kleidon** A.: Effects of tropical deforestation on surface energy balance partitioning in southeastern Amazonia estimated from maximum convective power. *Geophys. Rese. Lett.*, 46. <https://doi.org/10.1029/2018GL081625>, 2019.

Dhara, C., **Renner**, M., **Kleidon**, A.: Broad climatological variation of surface energy balance partitioning across land and ocean predicted from the maximum power limit. *Geophys. Rese. Lett.*, 43(14), 7686–7693. doi:10.1002/2016GL070323, 2016.

Hildebrandt, A., **Kleidon**, A., Bechmann, M.: A thermodynamic formulation of root water uptake. *Hydrology and Earth System Sciences*, 20(8), 3441–3454. doi:10.5194/hess-20-3441-2016, 2016.

**Kleidon**, A., & **Renner**, M.: A simple explanation for the sensitivity of the hydrologic cycle to surface temperature and solar radiation and its implications for global climate change. *Earth Syst. Dynam.*, 4(2), 455–465. doi:10.5194/esd-4-455-2013, 2013.

**Kleidon**, A., & **Renner**, M.: Thermodynamic limits of hydrologic cycling within the Earth system: concepts, estimates and implications. *Hydrol. Earth Syst. Sci.*, 17(7), 2873–2892. doi:10.5194/hess-17-2873-2013, 2013.

**Kleidon**, A., **Zehe**, E., **Ehret**, U., and Scherer, U.: Thermodynamics, maximum power, and the dynamics of preferential river flow structures at the continental scale, *Hydrol. Earth Syst. Sci.*, 17, 225–251, 2013.

**Kleidon**, A., **Renner**, M., & Porada, P.: Estimates of the climatological land surface energy and water balance derived from maximum convective power. *Hydrol. Earth Syst. Sci.*, 18, 2201–2218. doi:10.5194/hess-18-2201-2014, 2014.

**Kleidon**, A., Kravitz, B., **Renner**, M.: The hydrological sensitivity to global warming and solar geoengineering derived from thermodynamic constraints. *Geophys. Rese. Lett.*, 42(1), 138–144. doi:10.1002/2014GL062589, 2015.

**Kleidon**, A., **Renner**, M.: An explanation for the different climate sensitivities of land and ocean surfaces based on the diurnal cycle. *Earth Syst. Dynam.*, 8(3), 849–864. doi:10.5194/esd-8-849-2017, 2017.

**Kleidon**, A., **Renner**, M.: Diurnal land surface energy balance partitioning estimated from the thermodynamic limit of a cold heat engine, *Earth Syst. Dynam.*, 9, 1127–1140, <https://doi.org/10.5194/esd-9-1127-2018>, 2018.

**Renner**, M., **Hassler**, S. K., **Blume**, T., **Weiler**, M., Hildebrandt, A., Guderle, M., **Schymanski**, S. J., and **Kleidon**, A.: Dominant controls of transpiration along a hillslope transect inferred from ecohydrological measurements and thermodynamic limits, *Hydrol. Earth Syst. Sci.*, 20, 2063–2083, doi:10.5194/hess-20-2063-2016, 2016.

**Renner**, M., **Brenner**, C., **Mallick**, K., Wizemann, H.-D., Conte, L., Trebs, I., Wei, J., **Wulfmeyer**, V., **Schulz**, K., and **Kleidon**, A.: Using phase lags to evaluate model biases in simulating the diurnal cycle of evapotranspiration: a case study in Luxembourg, *Hydrol. Earth Syst. Sci.*, 23, 515–535, <https://doi.org/10.5194/hess-23-515-2019>, 2019.

#### 1.4.2.7 Phase II project D and phase I project A, PI L. Pfister, Team F. Fenicia, N. Martínez-Carreras, C. Hissler, L. Godol

**Fenicia**, F., Kavetski, D., Savenije, H.H.G., Clark, M. P., Schoups, G., **Pfister**, L., and J. Freer: Catchment properties, function, and conceptual model representation: is there a correspondence? *Hydrological Processes*, 28: 2451–2467, 2014.

**Martínez-Carreras**, N., Wetzel, C.E., Frentress, J., Ector L., McDonnell, J.J., Hoffmann, L. and **Pfister**, L.: Hydrological connectivity inferred from diatom transport through the riparian-stream system. *Hydrology and Earth System Sciences*, 19:3133–3151, 2015.

Klaus, J., Wetzel, C. E., **Martínez-Carreras**, N., Ector, L., and **Pfister**, L.: A tracer to bridge the scales: on the value of diatoms for tracing fast flow path connectivity from headwaters to meso-scale catchments, *Hydrol. Process.*, 29, 5275–5289, <https://doi.org/10.1002/hyp.10628>, 2015.

**Martínez-Carreras N., Hissler C., Gourdol L.,** Klaus J., Juilleret J., Iffly, J.F., **Pfister, L.:** Storage controls on the generation of double peak hydrographs in a forested headwater catchment. *J. of Hydrol.* 543: 255-269, 2016.

**Pfister, L., Martínez-Carreras, N., Hissler, C.,** Klaus, J., Carrer, G. E., Stewart, M. K., and McDonnell, J. J.: Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments, *Hydrological Processes*, 31, 1828-1845, 10.1002/hyp.11134, 2017.

**Pfister, L.,** Wetzel, C. E., Klaus, J., **Martínez-Carreras, N.,** Antonelli, M., Teuling, A. J., and McDonnell, J. J.: Terrestrial diatoms as tracers in catchment hydrology: a review, *Wiley Interdisciplinary Reviews-Water*, 4, 10.1002/wat2.1241, 2017.

Rodriguez, N. B., **Pfister, L., Zehe, E.,** and Klaus, J.: A comparison of catchment travel times and storage deduced from deuterium and tritium tracers using StorAge Selection functions, *Hydrology and Earth System Sciences*, 25, 401-428, 10.5194/hess-25-401-2021, 2021.

**Schwab M.P.,** Klaus J., **Pfister L.,** Weiler, M.: Diel fluctuations of viscosity-driven riparian inflow affect streamflow DOC concentration. *Biogeosciences* 15: 2177-2188, 2018.

Wrede, S., **Fenicia, F., Maretínez-Carreras, N.,** Juilleret, J., Krein, A., Savenije, H.H.J, Uhlenbrock, S., Kavetski, D. and **Pfister, L.:** Towards more systematic perceptual model development: A case study using 3 Luxembourgish catchments. *Hydrological Processes*, 29 (12): 2731-2750, 2015.

1.4.2.8 Phase II Project D and phase I project H, PI M. Weiler, T.Blume. Team. M Sprenger, L. Angermann, D. Demand

**Demand, D., Blume, T., and Weiler, M.:** Spatio-temporal relevance and controls of preferential flow at the landscape scale, *Hydrol. Earth Syst. Sci.*, 23, 4869–4889, <https://doi.org/10.5194/hess-23-4869-2019>, 2019.

**Demand, D.,** Selker, J.S., **Weiler, M.:** Influences of Macropores on Infiltration into Seasonally Frozen Soil, *Vadose Zone Journal*, 18:180147, doi:10.2136/vzj2018.08.0147, 2019.

**Angermann, L., Jackisch, C., Allroggen, N., Sprenger, M., Zehe, E., Tronicke, J., Weiler, M., and Blume, T.:** Form and function in hillslope hydrology: characterization of subsurface flow based on response observations. *Hydrol. Earth Syst. Sci.*, 21, 3727-3748, <https://doi.org/10.5194/hess-21-3727-2017>, 2017.

**Sprenger, M.,** Herbstritt, B. and **Weiler, M.:** Established methods and new opportunities for pore water stable isotope analysis. *Hydrological Processes*, DOI:10.1002/hyp.10643, 2015.

**Sprenger M.,** T. H. M. Volkmann, **T. Blume,** and **M. Weiler:** Estimating flow and transport parameters in the unsaturated zone with pore water stable isotopes. *Hydrol. Earth Syst. Sci.*, 19, 2617–2635, 2015.

**Sprenger M.,** Seeger S., **Blume T., Weiler M.:** Travel times in the vadose zone: variability in space and time. *Water Resour. Res.*, 52 (8),5727-5754, doi: 10.1002/2015WR018077, 2016.

**Sprenger, M.,** Gimbel, K., Leistert, H. and **Weiler, M.:** Illuminating hydrological processes at the soil-vegetation-atmosphere interface with water stable isotopes. *Reviews of Geophysics*, DOI: 10.1002/2015RG000515, 2016.

**Blume T.** and van Meerveld H.J.: From hillslope to stream: methods to investigate subsurface connectivity, *Wires Water*. 2015.

1.4.2.9 Phase II Project E and phase I Project S, PI U. Ehret, Team: M. Westhoff, M. Neuper, R. van Pruisen

Azmi, E., **Ehret, U.,** Weijs, S. V., Ruddell, B. L., and Perdigão, R. A. P.: Technical note: “Bit by bit”: a practical and general approach for evaluating model computational complexity vs. model performance, *Hydrol. Earth Syst. Sci.*, 25, 1103–1115, <https://doi.org/10.5194/hess-25-1103-2021>, 2021.

**Ehret, U.,** Gupta, H. V., Sivapalan, M., Weijs, S. V., **Schymanski, S. J.,** Blöschl, G., Gelfan, A. N., Harman, C., Kleidon, A., Bogaard, T. A., Wang, D., Wagener, T., Scherer, U., **Zehe, E.,** Bierkens, M. F. P., Di Baldassarre, G., Parajka, J., van Beek, L. P. H., van Griensven, A., **Westhoff, M. C.,** and Winsemius, H. C.: Advancing catchment hydrology to deal with predictions under change, *Hydrol. Earth Syst. Sci.*, 18, 649–671, <https://doi.org/10.5194/hess-18-649-2014>, 2014.

**Ehret, U., van Pruijssen, R.,** Bortoli, M., **Loritz, R.,** Azmi, E., and **Zehe, E.:** Adaptive clustering: reducing the computational costs of distributed (hydrological) modelling by exploiting time-variable similarity among model elements, *Hydrology and Earth System Sciences*, 24, 4389-4411, 10.5194/hess-24-4389-2020, 2020.

**Loritz, R.,** Hrachowitz, M., **Neuper, M., and Zehe, E.:** The role and value of distributed precipitation data in

- hydrological models, *Hydrol. Earth Syst. Sci.*, 25, 147–167, <https://doi.org/10.5194/hess-25-147-2021>, 2021.
- Neuper, M., and Ehret, U.:** Quantitative precipitation estimation with weather radar using a data- and information-based approach, *Hydrol. Earth Syst. Sci.*, 23, 3711-3733, [10.5194/hess-23-3711-2019](https://doi.org/10.5194/hess-23-3711-2019), 2019.
- Thiesen, S., Darscheid, P., and **Ehret, U.:** Identifying rainfall-runoff events in discharge time series: a data-driven method based on information theory, *Hydrol. Earth Syst. Sci.*, 23, 1015-1034, [10.5194/hess-23-1015-2019](https://doi.org/10.5194/hess-23-1015-2019), 2019.
- Westhoff, M. C., and Zehe, E.:** Maximum entropy production: can it be used to constrain conceptual hydrological models? *Hydrology And Earth System Sciences*, 17, 3141-3157, [10.5194/hess-17-3141-2013](https://doi.org/10.5194/hess-17-3141-2013), 2013.
- Westhoff, M. C., Zehe, E., and Schymanski, S. J.:** Importance of temporal variability for hydrological predictions based on the maximum entropy production principle, *Geophysical Research Letters*, 41, 67-73, [10.1002/2013gl058533](https://doi.org/10.1002/2013gl058533), 2014.
- Westhoff, M., Zehe, E., Archambeau, P., and Dewals, B.:** Does the Budyko curve reflect a maximum-power state of hydrological systems? A backward analysis, *Hydrology And Earth System Sciences*, 20, 479-486, [10.5194/hess-20-479-2016](https://doi.org/10.5194/hess-20-479-2016), 2016.
- Zehe, E., Ehret, U., Blume, T., Kleidon, A., Scherer, U., and Westhoff, M.:** A thermodynamic approach to link self-organization, preferential flow and rainfall-runoff behaviour, *Hydrology And Earth System Sciences*, 17, 4297-4322, [10.5194/hess-17-4297-2013](https://doi.org/10.5194/hess-17-4297-2013), 2013.
- Zehe, E., Ehret, U., Pfister, L., Blume, T., Schroeder, B., Westhoff, M., Jackisch, C., Schymanski, S. J., Weiler, M., Schulz, K., Allroggen, N., Tronicke, J., van Schaik, L., Dietrich, P., Scherer, U., Eccard, J., Wulfmeyer, V., and Kleidon, A.:** HESS Opinions: From response units to functional units: a thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments, *Hydrology And Earth System Sciences*, 18, 4635-4655, [10.5194/hess-18-4635-2014](https://doi.org/10.5194/hess-18-4635-2014), 2014.
- 1.4.2.10 Phase II Project F and phase I Project I, PI E. Zehe, CoPi Th. Blume Team: R. Loritz, C. Jackisch
- Glaser, B., **C. Jackisch**, L. Hopp, and J. Klaus.: How Meaningful are Plot-Scale Observations and Simulations of Preferential Flow for Catchment Models? *Vadose Zone Journal*, 18(1), 0, [doi:10.2136/vzj2018.08.0146](https://doi.org/10.2136/vzj2018.08.0146), 2019.
- Jackisch, C., L. Angermann, N. Allroggen, M. Sprenger, T. Blume, J. Tronicke, and E. Zehe.:** Form and function in hillslope hydrology: in situ imaging and characterization of flow-relevant structures, *Hydrol. Earth Syst. Sci.*, 21(7), 3749–3775, [doi:10.5194/hess-21-3749-2017](https://doi.org/10.5194/hess-21-3749-2017), 2017,
- Jackisch, C. and Zehe, E.;** Ecohydrological particle model based on representative domains, *Hydrol. Earth Syst. Sci.*, 22(7), 3639–3662, [doi:10.5194/hess-22-3639-2018](https://doi.org/10.5194/hess-22-3639-2018), 2018.
- Jackisch, C., Knoblauch, S., Blume, T., Zehe, E., and Hassler, S. K.:** Estimates of tree root water uptake from soil moisture profile dynamics, *Biogeosciences*, 17, 5787–5808, <https://doi.org/10.5194/bg-17-5787-2020>, 2020.
- Loritz, R., Kleidon, A., Jackisch, C., Westhoff, M., Ehret, U., Gupta, H. and Zehe, E.:** A topographic index explaining hydrological similarity by accounting for the joint controls of runoff formation, *Hydrol. Earth Syst. Sci.*, 23(9), 3807–3821, [doi:10.5194/hess-23-3807-2019](https://doi.org/10.5194/hess-23-3807-2019), 2019.
- Loritz, R., Hassler, S. K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J. and Zehe, E.:** Picturing and modeling catchments by representative hillslopes, *Hydrol. Earth Syst. Sci.*, 21(2), 1225–1249, [doi:10.5194/hess-21-1225-2017](https://doi.org/10.5194/hess-21-1225-2017), 2017.
- Loritz, R., Gupta, H., Jackisch, C., Westhoff, M., Kleidon, A., Ehret, U. and Zehe, E.:** On the dynamic nature of hydrological similarity, *Hydrol. Earth Syst. Sci.*, 22(7), 3663–3684, [doi:10.5194/hess-22-3663-2018](https://doi.org/10.5194/hess-22-3663-2018), 2018.
- Malicke, M., Hassler, S. K., Blume, T., Weiler, M., and Zehe, E.:** Soil moisture: variable in space but redundant in time, *Hydrology and Earth System Sciences*, 24, 2633-2653, [10.5194/hess-24-2633-2020](https://doi.org/10.5194/hess-24-2633-2020), 2020.
- Seibert, S. P., **C. Jackisch, U. Ehret, L. Pfister, and E. Zehe.;** Unravelling abiotic and biotic controls on the seasonal water balance using data-driven dimensionless diagnostics, *Hydrol. Earth Syst. Sci.*, 21(6), 2817–2841, [doi:10.5194/hess-21-2817-2017](https://doi.org/10.5194/hess-21-2817-2017), 2017.
- Sternagel, A., **R. Loritz, W. Wilcke, and E. Zehe.:** Simulating preferential soil water flow and tracer transport using the Lagrangian Soil Water and Solute Transport Model, *Hydrol. Earth Syst. Sci.*, 23(10), 4249–4267, [doi:10.5194/hess-23-4249-2019](https://doi.org/10.5194/hess-23-4249-2019), 2019.

**Zehe, E., Loritz, R., Jackisch, C., Westhoff, M., Kleidon, A., Blume, T., Hassler, S. K. and Savenije, H. H.:** Energy states of soil water – a thermodynamic perspective on soil water dynamics and storage-controlled streamflow generation in different landscapes, *Hydrol. Earth Syst. Sci.*, 23(2), 971–987, doi:10.5194/hess-23-971-2019, 2019.

**Zehe, E., and C. Jackisch:** A Lagrangian model for soil water dynamics during rainfall-driven conditions, *Hydrol. Earth Syst. Sci.*, 20(9), 3511–3526, doi:10.5194/hess-20-3511-2016, 2016.

1.4.2.11 Phase I and phase II Project G, PI T. Blume, M. Weiler, Team: S. Hassler, D. Beiter, E. Sohrt (geb. Lieder), N. Kaplan

**Hassler S.K., Weiler M., Blume T.:** Tree-, stand- and site-specific controls on landscape-scale patterns of transpiration. *Hydrol. Earth Syst. Sci.*, 22, 13–30, 2018, <https://doi.org/10.5194/hess-22-13-2018>

Hoek van Dijke, A. J., **Mallick, K.**, Teuling, A. J., **Schlerf, M.**, Machwitz, M., **Hassler, S. K., Blume, T.** & Herold, M. (2019). Does the Normalized Difference Vegetation Index explain spatial and temporal variability in sap velocity in temperate forest ecosystems? *Hydrology and Earth System Sciences*, 23(4), 2077-2091. <https://doi.org/10.5194/hess-23-2077-2019>

Casper, M. C., Mohajerani, H., **Hassler, S. K.**, Herdel, T., **Blume, T.:** Finding behavioral parameterization for a 1-D water balance model by multi-criteria evaluation. *Journal of Hydrology and Hydromechanics*, 67, 3, pp. 213-224. DOI: <http://doi.org/10.2478/johh-2019-0005>, 2019.

**Beiter, D., Weiler, M., and Blume, T.:** Characterising hillslope–stream connectivity with a joint event analysis of stream and groundwater levels, *Hydrol. Earth Syst. Sci.*, 24, 5713–5744, <https://doi.org/10.5194/hess-24-5713-2020>, 2020.

**Kaplan, N. H., Sohrt, E., Blume, T., Weiler, M.:** Monitoring ephemeral, intermittent and perennial streamflow: a dataset from 182 sites in the Attert catchment, Luxembourg. - *Earth System Science Data*, 11, pp. 1363-1374. DOI: <http://doi.org/10.5194/essd-11-1363-2019>, 2019.

**Kaplan, N. H., Blume, T., and Weiler, M.:** Predicting probabilities of streamflow intermittency across a temperate mesoscale catchment, *Hydrol. Earth Syst. Sci.*, 24, 5453–5472, <https://doi.org/10.5194/hess-24-5453-2020>, 2020.

**Blume T., van Meerveld I., Weiler M.:** The role of experimental work in hydrological sciences – insights from a community survey. *Hydrological Sciences Journal* 62 (3), 334-337, DOI: 10.1080/02626667.2016.1230675, 2017.

**Blume, T., van Meerveld, I., Weiler, M. (2018):** Incentives for field hydrology and data sharing: collaboration and compensation: reply to “A need for incentivizing field hydrology, especially in an era of open data”\*. - *Hydrological Sciences Journal - Journal des Sciences Hydrologiques*, 63, 8, pp. 1266-1268. DOI: <http://doi.org/10.1080/02626667.2018.1495839>

**Blume, T., van Meerveld, I., Weiler, M. (2018):** Why and when it is useful to publish and share inconclusive results and failures: reply to “Reporting negative results to stimulate experimental hydrology”. - *Hydrological Sciences Journal - Journal des Sciences Hydrologiques*, 63, 8, pp. 1273-1274. DOI: <http://doi.org/10.1080/02626667.2018.1493204>

Flavia Tauro, John Selker, Nick van de Giesen, Tommaso Abrate, Remko Uijlenhoet, Maurizio Porfiri, Salvatore Manfreda, Kelly Caylor, Tommaso Moramarco, Jerome Benveniste, Giuseppe Ciruolo, Lyndon Estes, Alessio Domeneghetti, Matthew T. Perks, Chiara Corbari, Ehsan Rabiei, Giovanni Ravazzani, Heye Bogena, Antoine Harfouche, Luca Brocca, Antonino Maltese, Andy Wickert, Angelica Tarpanelli, Stephen Good, Jose Manuel Lopez Alcala, Andrea Petroselli, Christophe Cudennec, **Theresa Blume**, Rolf Hut & Salvatore Grimaldi: Measurements and Observations in the XXI century (MOXXI): innovation and multi-disciplinarity to sense the hydrological cycle, *Hydrological Sciences Journal*, DOI: 10.1080/02626667.2017.1420191, 2018.

#### 1.4.3 Most important CAOS Conferences in the funding period

- AGU-CHAPMAN conference: Spatial organization and complex behavior of intermediate scale catchments – is there a connection? 23-26 September 2014, Luxembourg-city, Luxembourg,
- Tag der Hydrologie 2019: Information und Organisation in der hydrologischen Forschung und Praxis. 28- 29 March 2019, KIT Karlsruhe, Germany,
- EGU Leonardo Topical Conference 2019: *Global change, landscape ageing and the pulse of catchments*. 16-18 October 2019, University of Luxembourg, Luxembourg.

## 2 Final progress report

### 2.1 CAOS research questions, hypotheses and structure

#### 2.1.1 First funding period

The main objective of the CAOS project was to explore the link between spatial organisation and hydrological functioning of lower mesoscale catchments. While most hydrological landscapes are heterogeneous, this heterogeneity is at least partly spatially organised i.e. of systematic and structured nature. Based on the idea that these structured patterns of topography, soil, vegetation and biota are co-evolutionary fingerprints of past water, energy, and nutrient flow regimes (Dietrich and Perron, 2006; Phillips 2006), we hypothesized that this caused the development of functional units, which exhibit a distinct hydrological behaviour (**hypothesis H1**). More specifically we proposed that a catchment is composed of sub-hillslope scale elementary functional units EFU acting similarly with respect to the land-surface energy balance and related vertical water and energy fluxes during radiation driven conditions, and hillslope scale functional units (named lead topologies) exhibiting similar rainfall runoff behaviour. In consequence, we proposed that the typical structure and dynamic behaviour of functional units can be exemplarily characterised through replicated experiments and monitoring (**hypothesis H2**). Functional units were furthermore expected to be suitable building blocks for development of simplified and yet physically consistent hydrological models of adaptive complexity for lower mesoscale catchments (**hypothesis H3**). An adaptive model complexity was deemed to reflect the time variant complexity of hydrological dynamics and to save unnecessary computation time by avoiding redundant computations.

Study area of the CAOS project has been the Attert experimental basin in Luxembourg. The latter has been operated since 1994 by the former CRP-Gabriel Lippmann, today Luxembourg Institute of Technology (LIST) (e.g. Pfister et al., 2009; 2010). The Attert was selected because it is particularly well suited to explore how homogeneous geologies and landuse settings as well as different mixtures thereof control hydrological dynamics (e.g. Pfister et al., 2002). It consists of nine nested sub catchments that have homogenous and mixed geologies ranging from Devonian schist/slate, over marls to sand- and limestone, while the semi-oceanic climate is spatially rather uniform. To test the three overarching CAOS hypotheses we combined monitoring and experiments using technology from soil physics, geophysics, remote sensing, (tracer) hydrology and ecology. In line with H2 our scale oriented experimental design relied on spatial replicated and dense clustering of these different observation technologies, using a network of 46 automated sensor clusters as anchors. Experimental work was combined with ecological and hydrological model development and theory building, which particularly explored a thermodynamic perspective on hydrological dynamics (Zehe et al., 2014). Joint research in the first phase was structured into 10 subprojects listed in Table 1.

#### 2.1.2 Second funding period

Based on the most promising scientific achievements and lessons learnt in phase I we sharpened our main objective to the development of a holistic framework to explore how spatial organization alongside with spatial heterogeneity controls terrestrial water and energy cycles in intermediate scale catchments.

Table 1: Phase I subproject structure of the CAOS research unit and main PIs

Project	Title, PI and CoPI
A	<u>Interpreting dominant processes at the mesoscale with a top-down approach using flexible model structures for hypothesis testing</u> Dr. Fabrizio Fenicia, Dr. Laurent Pfister, Centre de Recherche Public - Gabriel Lippmann, Department 'Environment and Agro-Biotechnologies GEOSAT Research Unit'
B	<u>Dynamic patterns of land-surface characteristics related to water/energy fluxes at the hillslope and small catchment scale using remote sensing</u> Prof. Dr. Karsten Schulz Department of Geography, Ludwig-Maximilians-Universität München, Prof. Dr. Thomas Udelhoven, Remote Sensing and Geoinformatics Department, Trier University
C	<u>Quantitative Precipitation Estimation (QPE) by exploiting the potential of advanced radar observations and high-resolution data assimilation</u> Prof. Dr. Volker Wulfmeyer, Institute of Physics and Meteorology, University Hohenheim; Dr. Uwe Ehret, Institute of Water and River Basin Management, Karlsruhe Institute of Technology
D	<u>Thermodynamics and soil-vegetation-atmosphere transfer processes</u> Dr. Axel Kleidon, Biospheric Theory and Modelling Group, Max-Planck-Institute for Biogeochemistry, Jena,
F	<u>From structures to parameters: geophysical exploration at the hillslope and catchment scale</u> Prof. Dr. Jens Tronicke, Institute of Earth and Environmental Sciences, University of Potsdam, Dr. Peter Dietrich, Department of Monitoring and Exploration Technology, Helmholtz Centre of Environmental Research, Leipzig
G	<u>Linking internal pattern dynamics and integral responses - Identification of dominant controls with a strategic sampling design</u> Dr. Theresa Blume (PI), Section Hydrology, German Research Centre for Geosciences Potsdam, Prof. Dr. Markus Weiler (PI), Institute of Hydrology, University of Freiburg
H	<u>Residence times across scales: from plot to catchment scale</u> Prof. Dr. Markus Weiler (PI), Institute of Hydrology, University of Freiburg, Dr. Theresa Blume (PI), Section Hydrology, German Research Centre for Geosciences Potsdam
I	<u>From subsurface structures to functions and texture - linking virtual realities and experiments at the plot and hillslope scales</u> Prof. Dr. Erwin Zehe (PI), Institute of Water and River Basin Management, Karlsruhe Institute of Technology, Dr. Theresa Blume (CoPI), Section Hydrology, German Research Centre for Geosciences Potsdam,
J	<u>Biogenic soil structures: feedbacks between bioactivity and spatial heterogeneity of water storage and fluxes from plot to hillslope scale</u> GastProf. Dr. Boris Schröder (PI), Institute of Earth and Environmental Sciences, University of Potsdam, Prof. Dr. Jana Eccart (CoPI), Institute for Biochemistry and Biology, University of Potsdam
S	<u>An adaptive process based model framework for water-, energy- and mass cycles in lower mesoscale catchments</u> Prof. Dr. Erwin Zehe (PI), (CoPI) Dr. Uwe Ehret, Institute of Water and River Basin Management, Karlsruhe Institute of Technology

The guiding hypotheses H1 – H3 of phase I were integrated into a set of four broader research questions and related deliverables (Fig. 1), in order better reflect and integrate our joint research interests and to foster the most promising research activities of phase I:

- Q1: How to define, characterize and link hydrological functioning and underlying controls across a hierarchy of spatial and temporal scales, including their interactions with the soil, vegetation and the atmosphere?
- Q2: Does spatial organization in catchments imply the existence of a hierarchy of functional units, which act similar with respect to different hydrological functions and can we characterize their typical behavior in an exemplary manner?
- Q3: How to balance necessary complexity and falsifiability of catchment models to step beyond the input-output paradigm?
- Q4: What is the connection between thermodynamics, organizing principles, catchment structure and catchment functioning?

To better address these questions, we redefined the CAOS structure to a set of seven larger subprojects (Table 2). This new structure particularly improved cooperation across institutional boundaries within two major overarching work packages (Fig. 1).

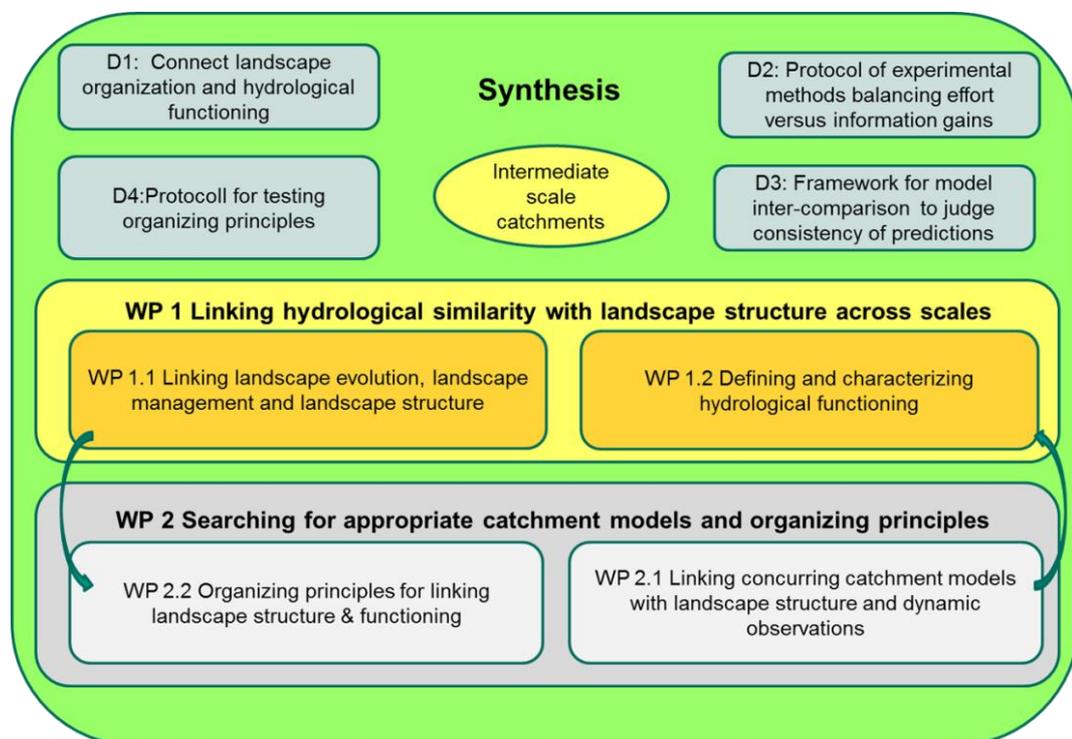


Figure 1: Phase II work package structure and main deliverables

Phase II project A investigated feedbacks between soils, biota, land management and hydrological processes at different spatiotemporal scales as continuation and extension of phase I project J (Biogenic soil structures: feedbacks between bioactivity and spatial heterogeneity of water storage and fluxes from plot to hillslope scale). This was done in close cooperation with phase II projects F and B as reflected in several joint publications (Schneider et al., 2018; van Schaik et al., 2016, Reck et al., 2018).

Phase II project B continued the successful development of geophysical methods, particularly of time lapse ground penetrating radar, to explore subsurface and flow process started in phase I project F (From structures to parameters: geophysical exploration at the hillslope and catchment scale, by Prof.

Tronicke). A new focus was to explore new ways to combine these geophysical methods with remote sensing methods, in cooperation with the group of Prof. Hinz (KIT), which joint the CAOS team at onset of Phase II. This project strongly cooperated with phase II projects A, F and G as reflected in several joint publications (Allrogen et al., 2017; 2016).

Table 2: Phase II subprojects and their relation to the overall work packages and deliverables shown in Fig. 1

Project	Title, PI and CoPI	Work package	Deliverable
A	<u>Feedbacks between soils, biota, land management and hydrological processes at different spatiotemporal scales</u> PI Prof. Schröder-Esselbach, Dr. van Schaik	1.1 1.2 2.1	D1 D3 D4
B	<u>Non-invasive geophysical and remote sensing methods to map and characterize relevant structures and processes</u> PI Prof. Tronicke, Prof Hinz	1.1 1.2	D2
C	<u>Understanding and characterizing land surface-atmosphere exchange and feedbacks</u> PI Prof. Kleidon, Prof. Wulfmeyer, Prof. Schulz, Dr. Mallick, Dr. Schlerf	1.1 2.2 1.2 2.1	D2 D4 D3 D1
D	<u>Spatio-temporal dynamics of water storage, mixing and release</u> PI Prof. Pfister, Dr. Hissler, Dr. Matgen, Dr. Blume, Prof. Weiler	1.2 2.1 1.1 2.2	D1 D3
E	<u>Towards consistent predictions of water and energy cycles in intermediate scale catchments</u> PI Dr. Ehret, Prof. Zehe	2.2 1.2	D3 D1
F	<u>Linking landscape structure, biota and hydrological functioning in an thermodynamic optimality context</u> PI Prof. Zehe	2.2 2.1 1.2	D4 D3 D1
G	<u>Hydrological connectivity and its controls on hillslope- and catchment scale stream flow generation</u> PI Prof. Weiler, Dr. Blume	1.2 2.2 1.1	D1 D4 D2

Phase II projects C and D were newly shaped to explore and define hydrological functioning separated as “*Land surface-atmosphere exchange and feedbacks*” during radiation driven conditions and “*Water storage, mixing and streamflow release*” during rainfall driven conditions. Project D is a continuation of phase I project H (Residence times across scales: from plot to catchment scale). Project C was newly designed to foster and advance research activities of 1) phase I project C (Quantitative Precipitation Estimation (QPE) by exploiting the potential of advanced radar observations and high-resolution data assimilation, group Wulfmeyer), 2) phase I project D (Thermodynamics and soil-vegetation-atmosphere transfer processes) and 3) phase I project B (Dynamic patterns of land-surface characteristics related to water/energy fluxes at the hillslope and small catchment scale using remote sensing).

Phase II project E largely continued the development and testing of new model concepts of adaptive complexity started in Phase I project S (An adaptive process-based model framework for water-, energy- and mass cycles in lower mesoscale catchments; Ehret et al., 2020; Loritz et al., 2021) as well as to advance the radar based QPE started within phase Project C (Neuper and Ehret, 2019). This was combined with a new perspective on data and model complexity drawing form information theoretic

concepts and measures in a close cooperation with phase II projects F and D (Neuper and Ehret, 2019; Darscheid et al., 2018, Ehret et al., 2020).

Phase II project F explored an alternative thermodynamic perspective on hydrological dynamics, in close cooperation with projects E and C (group Kleidon), as reflected joint publications (Zehe et al. 2019; Loritz et al. 2017; 2018; 2019). Strong emphasis was also on the advancement of research activities from phase I project I (From subsurface structures to functions and texture linking virtual realities and experiments at the plot and hillslope scales). This included new ways for in-situ imaging of subsurface structure and rapid subsurface flow (Jackisch et al., 2017) in close cooperation with project B, and the development of Lagrangian models for simulating soil water dynamics in heterogeneous soils (Zehe and Jackisch, 2016; Jackisch and Zehe, 2018; Sternagel et al., 2019).

Phase II project G focused on a detailed experimental characterisation of surface and subsurface connectivity using novel experimental approaches. This effort was a continuation and advancement of research conducted in phase I project G (Linking internal pattern dynamics and integral responses - Identification of dominant controls with a strategic sampling design).

In the following sections, we will detail on our joint experimental, modelling and theory building activities. We then close with a summary of our scientific achievements and failures, wired around the research questions as well as the central CAOS hypothesis postulating the existence of functional units of similar hydrological behaviour.

## **2.2 Project development**

### **2.2.1 Experimental characterisation of landscape structure and functioning**

Experimental efforts in phase I were focused on characterising the plot, hillslope and headwater scales. A key challenge was to balance the need for an exhaustive characterization of structure and processes within pre-selected EFU and Lead topologies (LT) candidates with the need to conduct replicates to detect typical functional and structural characteristics. Within phase II we extended our experimental efforts to the entire Attert catchment.

#### 2.2.1.1 The CAOS monitoring network to characterise landscape structure and functional units

##### 2.2.1.1.1 Conceptual framework and selection of candidate functional units

The conceptual starting point to set up the CAOS observation network was the well-known fact that any flux is the product of a potential gradient and the inverse of a (control volume) resistance. While the gradient specifies the driving force or energy difference, the latter characterises the dissipative energy losses along the flow path (Zehe et al. 2014). As sufficient criterion for functional similarity, we hence defined similarity of those landscape characteristics, which control both the gradient, and the resistance of the flux of interest. In consequence we related the driving gradients and resistance terms of relevant hydrological fluxes to observable system characteristics (Table 3), special emphasis was on relevant connected preferential pathways, because these reduce dissipative losses and thus the flow resistance (Zehe et al., 2013).

Table 3: Gradients and resistances determining fluxes/storage of water and energy as well as their landscape controls, special emphasis is also on the influence of connected network like structures, which reduce resistances. Note that resistance are rather control volume than material properties here.

Processes	Gradient	Landscape control	Resistance	Network like structure
<b>Energy exchange and storage</b>				
Transpiration	Vapor pressure canopy-atmosphere	Canopy albedo and temperature Aspect and slope Air vapor pressure Soil water potential Wind speed	Canopy and boundary layer resistances, root resistance, plant physiology	Canopy structure, Leaf area index (LAI), root network topology
Evaporation	Vapor pressure soil -atmosphere	Soil albedo and temperature Aspect and slope Soil water content & soil water retention curve Wind speed	Inverse of soil hydraulic conductivity Boundary layer resistance	Pore network
Sensible heat flux	Temperature surface-atmosphere	Soil albedo and temperature Aspect and slope Surface roughness Air temperature Wind speed	Turbulent/laminar boundary layer resistance	
Soil heat flux	Soil temperature	Soil albedo and temperature Aspect and slope Heat capacity Soil water content	Inverse of soil thermal conductivity content	Advective heat flux
<b>Water storage and drainage</b>				
Surface runoff	Overland flow depth	Surface topography & permeability	Surface roughness (incl. plants and debris),	Rill network topology & spec. flow resistance
Infiltration	Soil hydraulic potential	Soil water retention curve, soil water content, depth to ground water	Inverse of hydraulic conductivity, soil water content	Macropore network topology & spec. flow resistance
Root water uptake	Water potential soil-root	Rooting depth Fine root distribution Canopy water demand Soil water content Depth to groundwater	Root system resistance Inverse of hydraulic conductivity	Root network Macropore network
Subsurface storm flow	Hydraulic head	Bedrock topography & permeability	Inverse of hydraulic conductivity, soil water content	Lateral pipe network & spec. flow resistance
Ground water flow	Piezometric head	Aquitard topography, specific storage coefficient	Inverse of hydraulic permeability	Fracture network topology & spec. flow resistance

From Table 3 we identified those landscape factors, which guided delineation of candidate functional units and their experimental characteristics (Table 4).

Table 4: Hierarchy of proposed functional classification scheme guiding the CAOS monitoring network

<b>Hierarchy level</b>	<b>Similarity</b>	<b>Descriptors</b>	<b>Preferential flow path</b>	<b>Dominance</b>
Hydro-Geomorphic Unit (catchment scale)	Base flow, runoff, active ground water storage	Geological setting, hydrogeology, geomorphology	River network	Permanent, long term
Lateral Topologies (hillslope scale)	Rainfall runoff transformation, free water storage	Potential energy differences: surface & bedrock topography, catena, aspect	Vertical macropore, lateral pipe or rill network	Rainfall driven conditions
Elementary Functional Unit EFU (field scale)	Land surface energy exchange/ET, capillary soil water supply	Slope position & aspect, landuse, soil type	Vegetation, Landuse	Radiation driven conditions

Instrumentation started in and strongly focused on the Colpach sub catchment, a major tributary of the Atert, which is entirely located on Devonian schist/slate. Based on a joint analysis of land use, available pedological and geomorphic properties data, we selected two candidate lead topologies:

- Short hillslopes with small riparian zone and deciduous forest with shallow Cambisols. This lead topology is characteristic for forested headwaters of the northern part of the Colpach catchment. We distinguished northern and southern aspect, as different soil structures could have developed in response to the different energy input, biomass production and litter fall. We hypothesized that subsurface storm flow dominates stream flow generation during rainfall events. Bedrock topography of the schist interface was deemed the most important time invariant determinant for the driving lateral gradient. The thickness of the weathered schist layer on top of the bedrock together with its porosity and lateral permeability determine the maximum storage volume as well as the lateral flow resistance for subsurface storm flow.
- Hillslopes with small riparian zone and pasture with shallow Cambisols. This LT is characteristic for the headwaters of the southern Colpach catchment and thus characterized by convergent flow paths and temporary wetlands in the near-stream and source areas. Bedrock topography was expected to be of importance here.

In total, we instrumented four replicates of the first LT and two of the second.

#### 2.2.1.1.2 Selection and instrumentation of candidate elementary functional units

We then characterized the structure and dynamic behaviour of EFUs and LTs by combining different observation technologies in a scale oriented hierarchical manner. A strategic cornerstone was the permanent monitoring using automated sensor clusters designed and build within phase I project G (Fig. 2a). These sensor clusters provided temporarily highly resolved information on fluxes and state variables controlling potential gradients and subsurface flow resistance. They consisted of one to five rain gauges; climate sensors for air temperature, relative humidity, wind speed and global radiation; ten sensors measuring soil moisture, electric conductivity and soil temperature in 3 depth profiles; three matric

potential sensors next to the soil moisture sensors; four water level sensors incl. temperature and EC to observe groundwater and stream water level fluctuations, and finally five sap flow sensors to estimate transpiration fluxes (Fig. 2b). For more details, please refer to the project G final report.

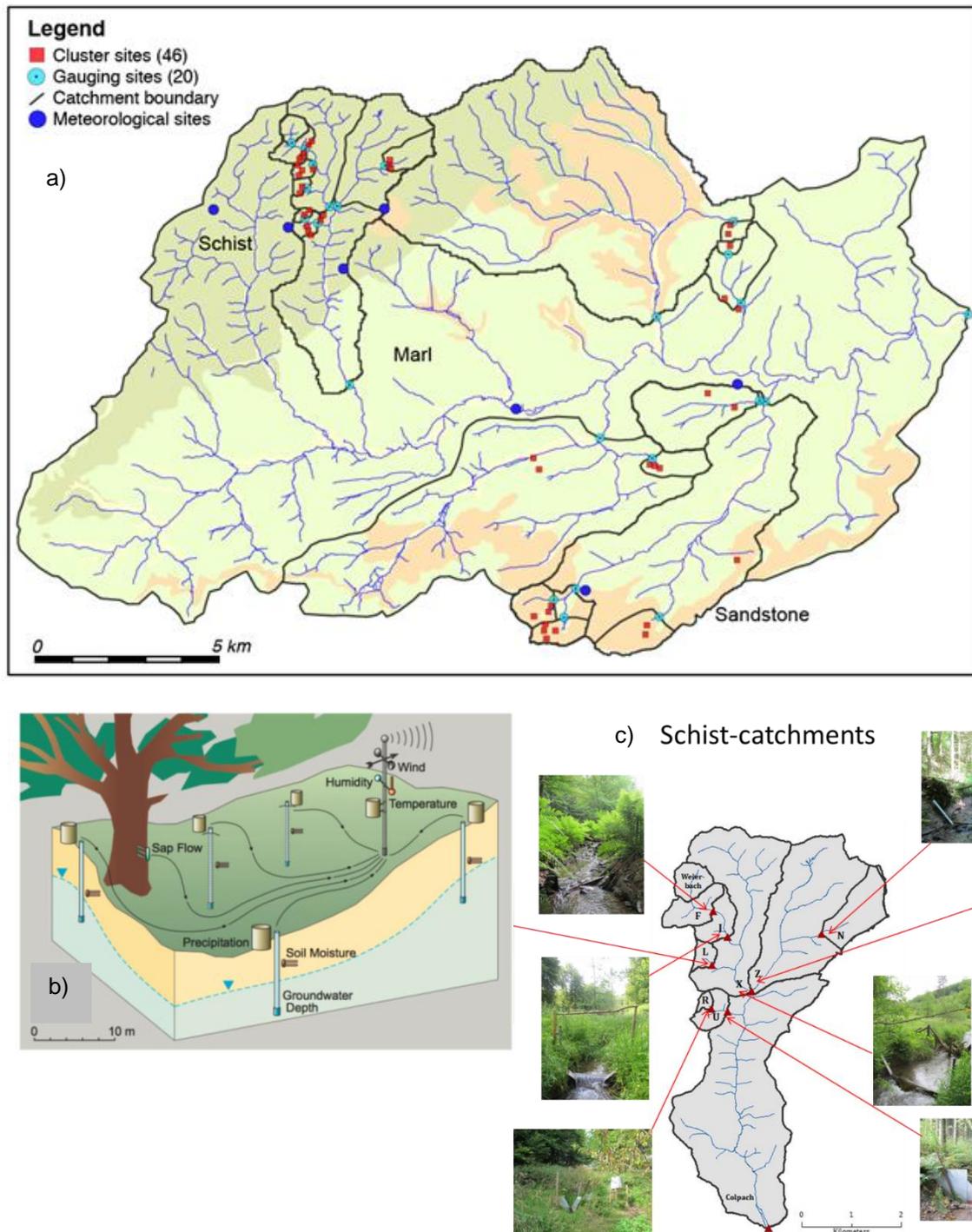


Figure 2: CAOS-monitoring network in the Attert basin (panel a). Layout of a sensor cluster (panel b). Gauged headwaters in the Colpach catchment (panel c).

Instrumented candidate EFUs in the schist area amounted to six on north facing slopes and 10 on south facing slopes, 7 units were situated close to a stream, and we included 16 forest and seven pasture sites. We expected that this design provided sufficient number of members of each candidate EFU class to enable characterization of similarities and differences. While most of the experimental effort was

focused on the schist area (23 sensor clusters), the sandstone and marl areas were also instrumented with 12 and 11 clusters, respectively. Coordination and most of the effort of installing the sensor clusters lay with phase I project G and H, but all working groups contributed to this installation. Especially in the Schist, the clusters were arranged within several headwater catchments, which were equipped with eight new stream gauges.

#### 2.2.1.1.3 Synoptic assessment of landscape structure and similarity

Phase I project B collected multiple satellite images for the past decades for the Attert catchment, including the high spatial resolution (30 m) Landsat TM/ETM data, medium spatial resolution (250 m - 1 km) MODIS data and 1 km SPOT-VGT data. Using a maximum likelihood classification, these data were transformed into reliable land cover maps, which allow analysis of landuse change dynamics within the last 40 years. Sun and Schulz (2015a) analyzed the potential of satellite Thermal Infra-Red data (TIR) for improving land cover classification in Phase II project C. Ground-truth reference data and TIR scenes over the Attert catchment, Luxembourg, were used to train and validate two classification methods, k-NN clustering and the random forest algorithm. It could be shown, that a time series of TIR images alone produces land cover maps with accuracies comparable to those derived from visible (VIS) and near infrared (NIR) spectral bands, while the combination of the TIR band with the VIS/NIR bands further increased the classification accuracy. Müller et al. (2014) derived furthermore a classification into landscape units as candidate EFUs based on ASTER thermal infrared remote sensing data. This revealed a distinct relationship between the first two principle components of the land-surface temperature time series with landuse/land cover and geology.

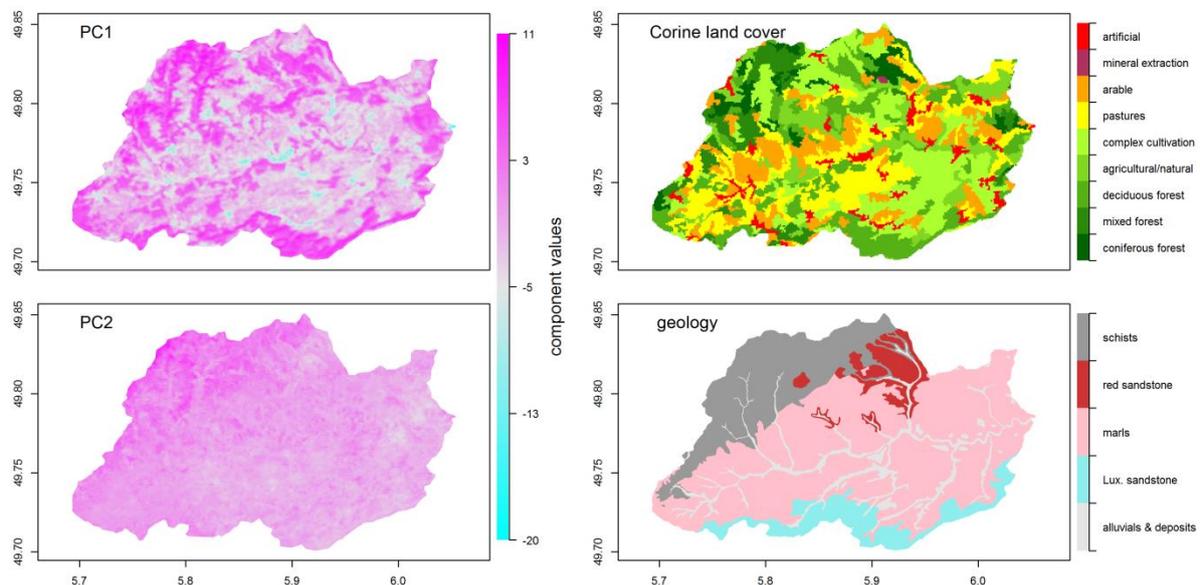


Figure 3: The first and second component of the principle component analysis (PCA) for the ASTER time series data (left), for comparison illustrations of Corine land cover and geology data (middle) of the Attert catchment.

Spatially distributed information on soil hydraulic properties is critical for many environmental research issues including, in the case of hydrological modelling, e.g. the understanding of water and solute transport processes below the surface. Commonly, hydraulic properties are estimated via pedo-transfer functions using soil texture data in order to avoid cost-intensive measurements of the hydraulic

parameters in the laboratory. Unfortunately, high-resolution spatially distributed soil texture data is rarely available. To fill this gap, Müller et al. (2016) developed a method to derive high-resolution (15 m) spatial topsoil patterns from a time series of satellite TIR observations within Phase I project B. Principle components of the TIR images, which reveal the most dominant thermal patterns within the landscape, were related to 212 fractional soil texture samples using a multiple linear regression framework. The proposed method yields soil texture estimates with comparable or even higher accuracy than reported in other recent studies.

Jackisch (2015) compiled a synoptic landscape analysis within phase II project F to search for distinct structural similarities between the three main geological settings. This include distributed soil sampling and measurements of saturates hydraulic conductivity,  $K_{sat}$ , at a range of locations (Fig. 4)

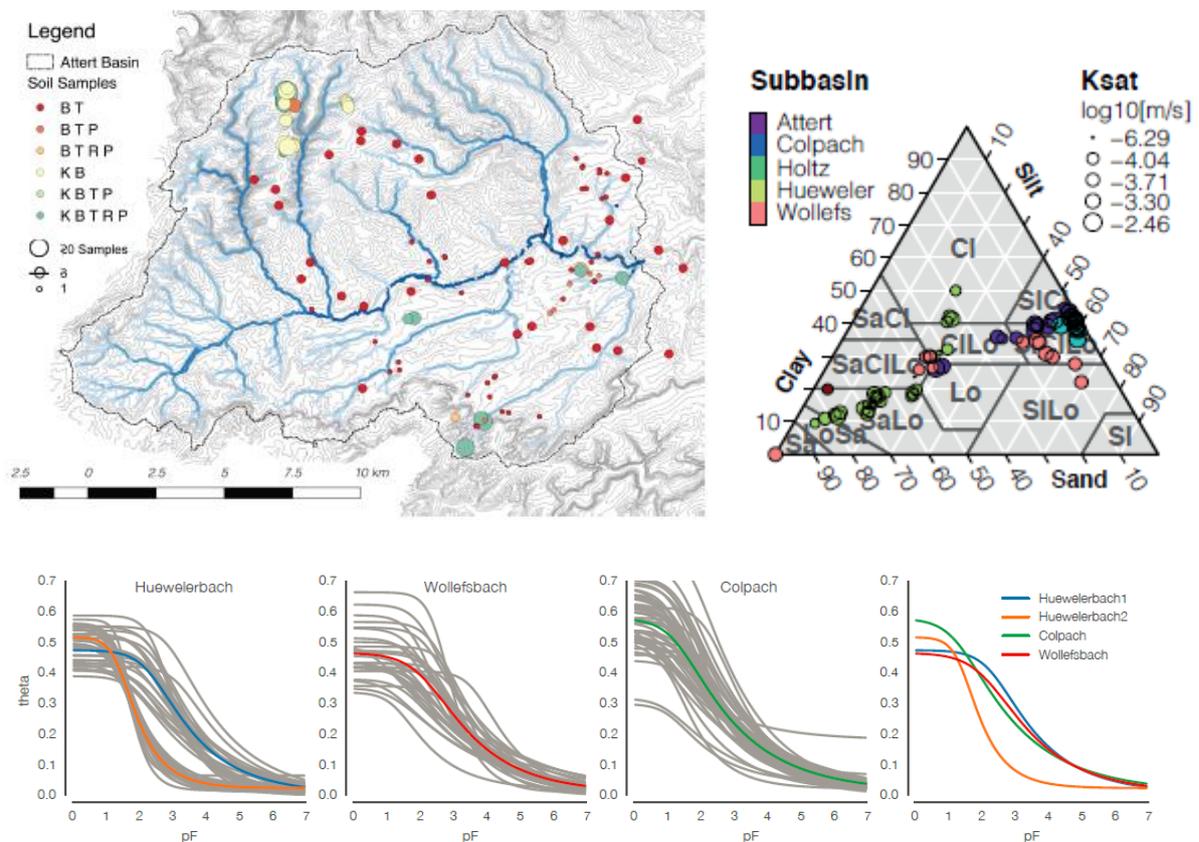


Figure 4: Location of soil core samples in the Atert basin (upper left panel). Colours indicate the respective analysis set, size give the number of samples at each location. B: Bulk Density (382 samples), T: Texture (wet sieving and sedimentation) (265) samples), P: pH (128 samples), R: Retention Curve (HYPROP apparatus) (120 samples), K: Ksat (Ksat apparatus) (223 samples). Upper right panel: texture and saturated hydraulic conductivity (size and taint of dot) of soil core samples in the Atert basin. Lower panels show retention functions for the Huewelerbach (Sandstone and Marls), Wollefsbach (Marls) and Colpach (Schist) as well as the corresponding effective curves.

The additional comparison of land use/ land cover, topographic signatures (topographic wetness index TWI (Beven and Kirkby, 1979), flow gradient, solar radiation index) and remote sensing data (Aster NIR, LAND-SAT normalized vegetation index NDVI, and normalized water index NWI), revealed systematically a strong similarity of the sandstone and schist area, while the marl setting appeared distinctly dissimilar (Fig. 5).

A synoptic analysis of the available rainfall runoff data set revealed however similarity among Marl and Schist catchments, while sandstone catchments performed distinctly different (compare section 2.2.1.3).

### 2.2.1.2 Characterising sub structure and preferential flow across scales and geological settings

To characterise the interplay of vertical and lateral subsurface structures and preferential flow processes we combined distributed tracer tests, with a soil survey, distributed sampling of earthworms and related dye tracing and ground penetrating radar as detailed in sections 2.2.1.2.1 and 2.2.1.2.3.

#### 2.2.1.2.1 Combined tracer tests, soil survey and earthworm sampling

Within phase I projects J and I we investigated the role of preferential flow paths by means of dye staining of flow paths and bromide tracer tests carried out at different irrigation intensities in all geological settings. This revealed a dominance of rapid flow with tracer travel depth of up to 2m within 1 day (Fig. 6).

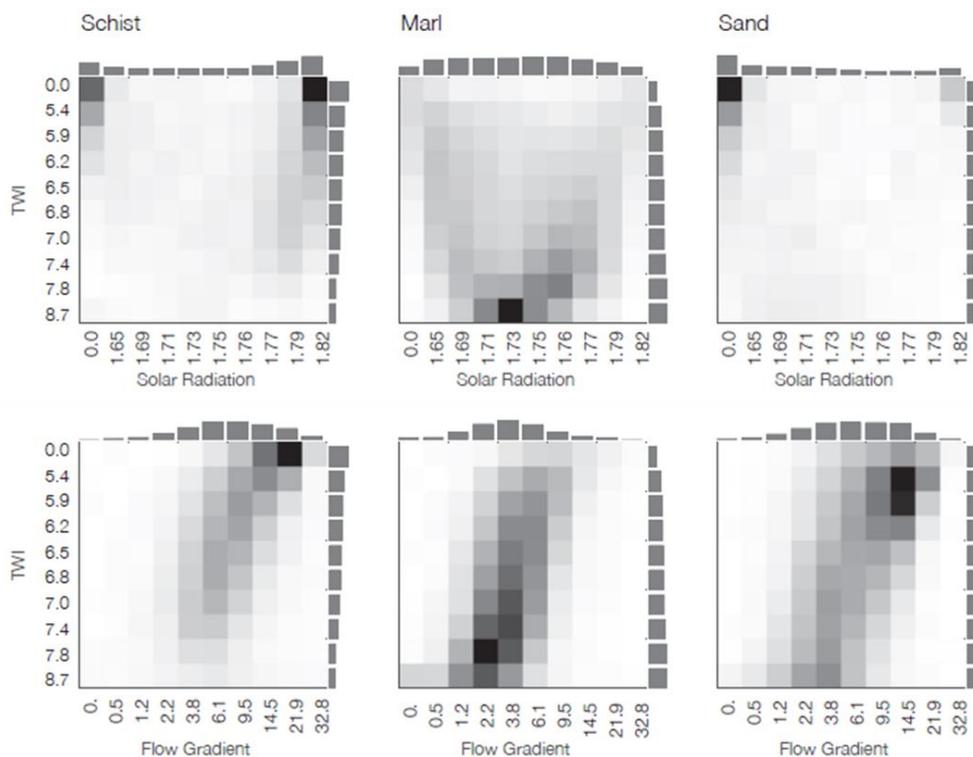


Figure 5: Bivariate distributions of TWI against a) solar radiation index, and b) flow gradient. Marginals give the overall distribution of the respective attribute in the geology class. Grid with black for maximum number of cells to belong to both respective classes. White refers to the minimum. Source: Jackisch (2015 PhD Thesis).

Moreover, we established a workflow to derive the soil water retention and unsaturated hydraulic conductivity curves using a large set of undisturbed soil cores from selected sites within all geological settings. This was combined with soil augers, constant head permeameter measurements and detailed soil type mapping in soil pits and dye staining (Jackisch et al., 2017). Overall, more than 120 soil cores were analysed for their retention properties and their saturated hydrologic conductivity (see Fig. 4). These data revealed that rapid flow in the silty soils in the Devonian slates operates within a well-connected network of inter-aggregate pores (Jackisch et al. 2017). Using this set of distributed retention

experiments Jackisch et al. (2017) derived macroscale effective soil water retention curves (Fig. 4) to sustain modelling efforts using representative hillslopes (Loritz et al., 2017) as detailed in section 2.2.2.1. These distributed dye tracer experiments went along with an ecological survey of the abundance and number of individuals of soil ecosystem engineers creating vertical and lateral preferential flow paths (phase I project J, phase II project A). We sampled different earthworm species and small rodents in a randomly stratified design at 117 plots, including the cluster sites if possible, considering the gradients of different habitat factors covering the entire catchment (Fig. 7). These data were used to successfully train species distribution models (Palm et al., 2013; Schneider et al., 2018). Further research within the second funding phase revealed a distinct relation between the abundance of worm up to 16 different worm species and hydrologically relevant biopores (van Schaik et al., 2014; 2016) as well as a clear time variance of hydrologically relevant active macropores (compare section 2.2.2).

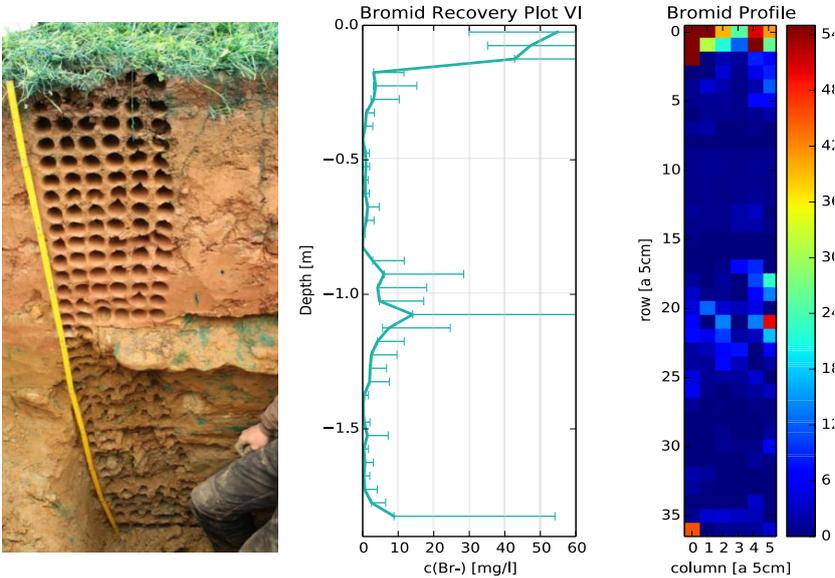


Figure 6: Example of bromide profile collected in the Colpach, sampling technique and bromide recovery.

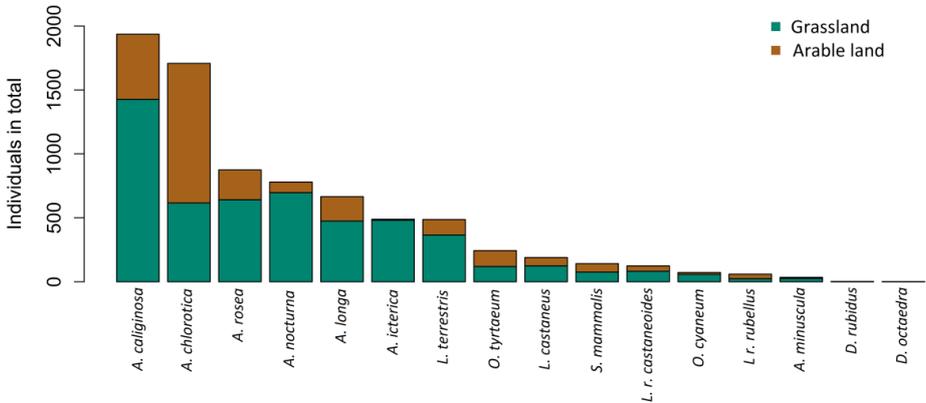


Figure 7: Total number of individuals of the 16 different earthworm species found throughout the field campaigns, marked green for meadow and brown for arable land. Source: Schneider et al., (2018).

These data and insights yielded valuable information for setting up a) the Lagrangian model echorRD, developed in phase II project F (Reck et al., 2018) to predict plot scale preferential transport (see section 2.2.2.3.1), as well as b) the representative hillslope model for the Wollefsbach catchment (Loritz et al., 2017). For more details please refer to final reports of phase II project A (which covers related research from phase I project J) and phase II project F (which summarizes the findings of phase I project I).

#### 2.2.1.2.2 In situ geophysical imaging of subsurface structures and flow processes

Furthermore, phase I project F and later on phase II project B evaluated different geophysical techniques for imaging and detection of subsurface structures and later on for in-situ monitoring and quantification of vertical and lateral subsurface water flow (Allroggen et al, 2015). Seismic techniques were unfeasible for the field sites in the Attert, because of the specific ground surface conditions, prevalent surface vegetation, and the shallowness of the targets. Electrical resistivity tomography (ERT) and particularly ground penetrating radar (GPR) turned out to be the most valuable techniques. The combination of ERT and GPR was proven to provide important information on depth to bedrock as well the depth of the weathered schist layer (Fig. 8). Both signatures were used to evaluate the consistency of the first-guess lead topologies and to estimate the downslope extend of EFUs (Zehe et al., 2014). Depth to bedrock and bedrock topography was also a key information to successfully setup a representative hillslope model for the Colpach catchment, which allowed a successful simulation runoff generation and the underlying fill and spill mechanism (Loritz et al., 2017).

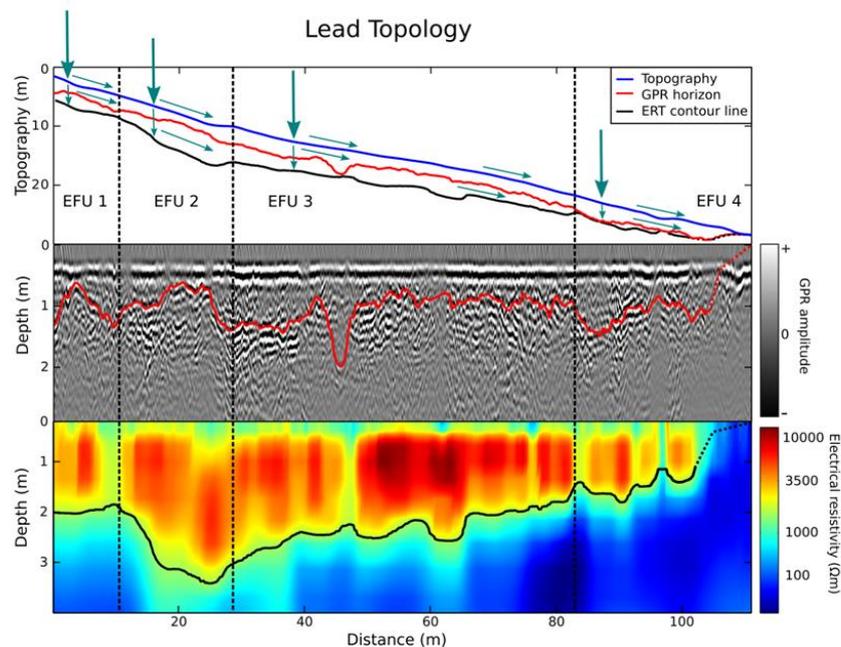


Figure 8: Geophysical approach to assess subsurface structures: ERT (bottom) and GPR (middle) surveys including an interpretation regarding subsurface layering and elementary functional units. Source: Zehe et al. (2014).

Moreover, a joint rainfall experiment, carried out by phase I projects and H and I, F, G, corroborated the feasibility of time laps GPR in combination with TDR profiling to visualise and quantify vertical and lateral subsurface flows (Allroggen et al, 2016; Angermann et al., 2017; Jackisch et al.; 2017). In line with our perception, the experiment revealed significant lateral flow in the subsurface geogenic structures, especially in the depths below 1 m (Fig. 9b). These structures were deemed to develop a spatially

connected subsurface network, allowing an efficient but relatively slow drainage of the long hillslopes in the Colpach catchment. This network was proposed to explain the frequently observed bimodal runoff responses in this area (Angermann et al., 2017, section 2.2.1.3.4). For more details on this experiment please refer to Angermann et al. (2017), Jackisch et al. (2017) or the final report of phase II project D (group Weiler phase I activities). Most recent findings corroborated that 4-D time lapse GPR during rainfall experiments became possible thanks to a newly designed and constructed irrigation device. This device allows for the first time in-situ monitoring of infiltration during irrigation and related quantification of soil moisture changes from relative reflectivity and suitable petrophysical relations (Fig. 10 and 11).

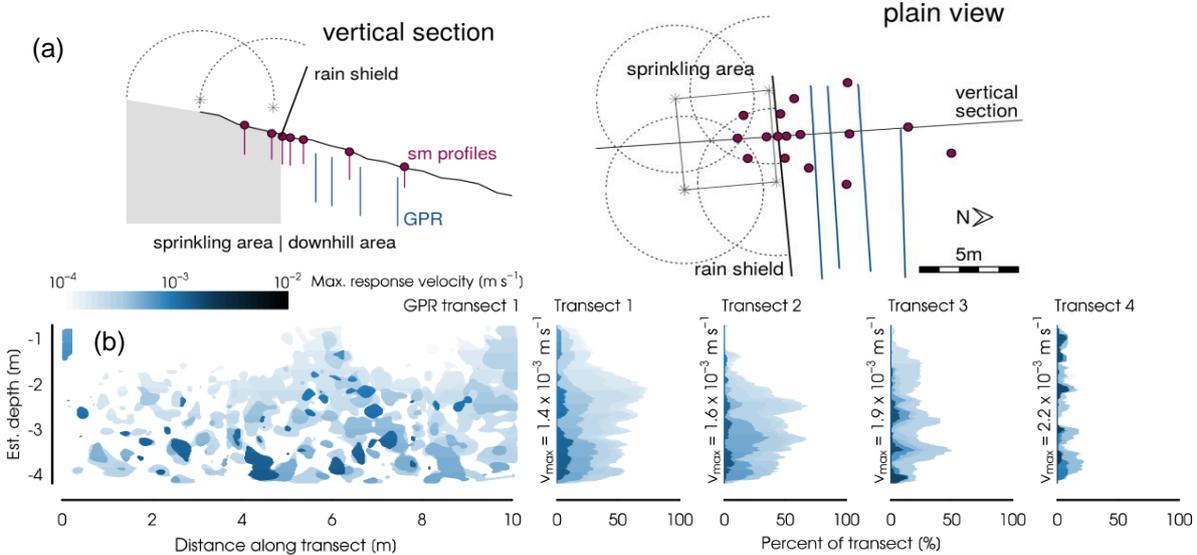


Figure 9: Experimental setup (a). Rapid lateral preferential flow at the hillslope scale: maximum response velocities estimated over the GPR transects. Source: Angermann et al. (2017).

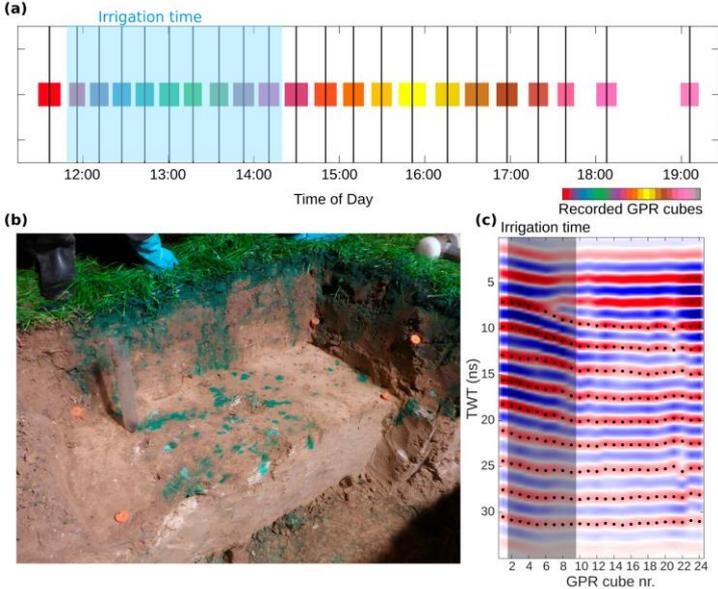


Figure 10: Field experiment in the Weiherbach catchment: (Germany) a) sequence of 4D GPR recording, b) dye staining of flow patterns and c) time sequence of single GPR trace. (Results to be published in Project B, Group Tronicke)

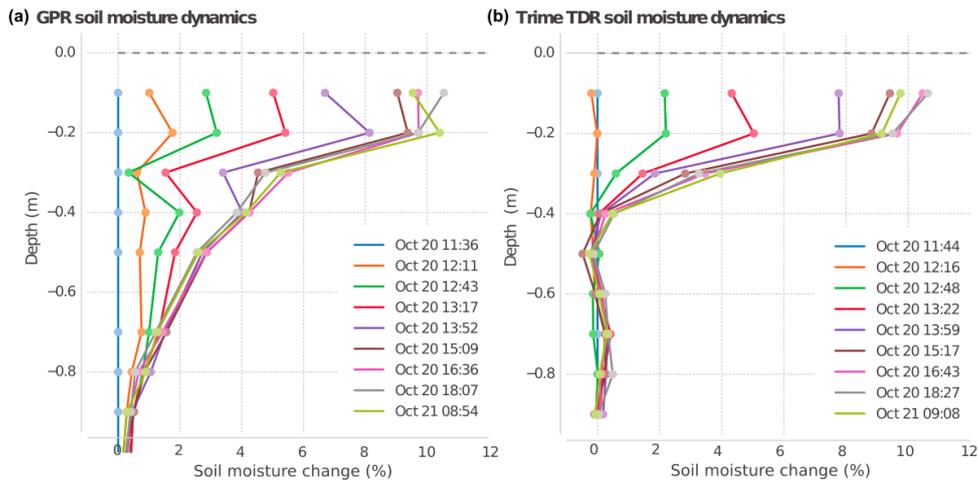


Figure 11: Soil moisture changes derived from a) reflectivity differences in GPR and (b) TDR measurements using PICO-IPH/T3 probe. (Results to be published in Project B, Group Tronicke)

For more details please refer to final reports of phase II projects B (group Tronicke, which also covers phase I project F), as well as to phase II projects A and F final reports for related co-operations.

Phase II Project G and Project F carried out an additional tracer experiment in the Weierbach catchment to further investigate subsurface flow. Tracer injection occurred via an upslope trench, which was filled with a constant level of salt solution. Thus, infiltration only occurred in this trench and the tracer propagated from there vertically and laterally downslope. This tracer movement was monitored through piezometers (measuring water level and electric conductivity), ERT- and GPR transects (Fig 12).

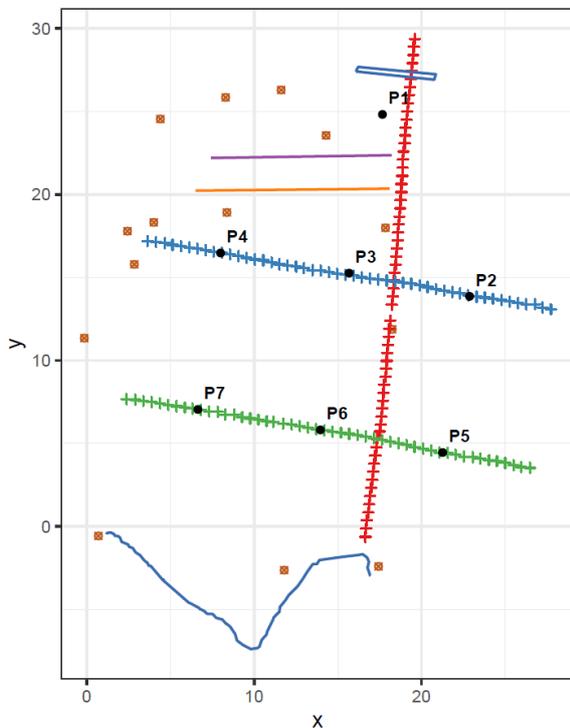


Figure 12. Layout of the three ERT transects (red = long transect, blue and green = cross- transects). The two GPR transects (purple and orange lines – note that there is a third one parallel to the orange line just downslope of the trees), the seven monitoring wells (P1 – P7), the trench (blue rectangle at top), the stream bank (blue line at bottom) and the bigger trees (brown points).

Figure 13 shows the responses in water level and EC, captured by the piezometers, where the blue background depicts the time of infiltration on each day. P1 shows the strongest response in both, while the next piezometer downslope (P3) barely reacts and the most downslope P6 reacts more strongly. The two piezometers at the side P2 and P4 don't show any reaction, meaning that the angle of the infiltration front from the trench did not open up that widely. In contrast, the two outer piezometers at the

foot-slope P5 and P7 show responses in water table and EC. These observations suggest that the infiltration pattern around the trench is mostly vertical until it reaches a confining layer that redirects the front in direction downslope. The front dives below the middle transect but is squeezed out when it

reaches the river bank, where the thickness of the soil layers becomes relatively small. When comparing the responses between the two days, one can observe a similar response in water table (except Piezo5) but a very different response in EC, with higher maximum values on day 2 but still far below the maximum values of Piezo1. The overall lower values in EC show that either the front is being diluted on its way through the soil, or that stratification of EC concentration occurs due to different density (the sensor measures on a particular depth), or a combination of both. The fact that the second day shows higher EC responses than the first points to accumulation of tracer in the subsurface that is mobilized on the second day, increasing EC.

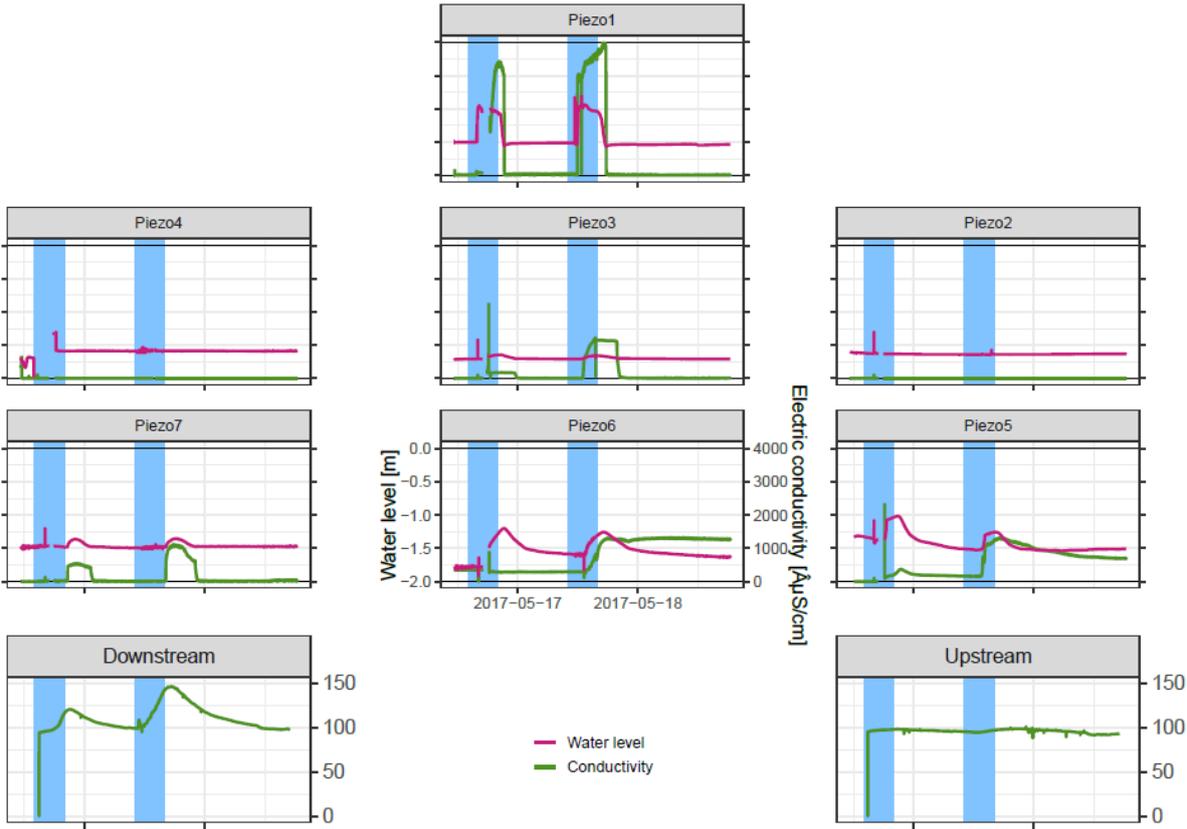


Figure 13: Water level and EC in piezometers 1-7 plus EC upstream and downstream of site. The position of each panel represents the location on the hillslope. The blue background marks times of trench infiltration, purple lines corresponds to the water level (on primary y-axis), green lines to the EC (secondary y-axis). Note that EC scale differs between piezometers (upper three rows) and stream (bottom row). (Results to be published in Project G).

Time-lapse ERT revealed at one hand a clear vertical and downslope migration of the tracer plume, as reflected in relative changes in resistivity (Fig 14). For snapshots after the reference, ratios are better indicators for changes than the actual model. Observations revealed however also an unrealistic increase in resistivity for higher depths.

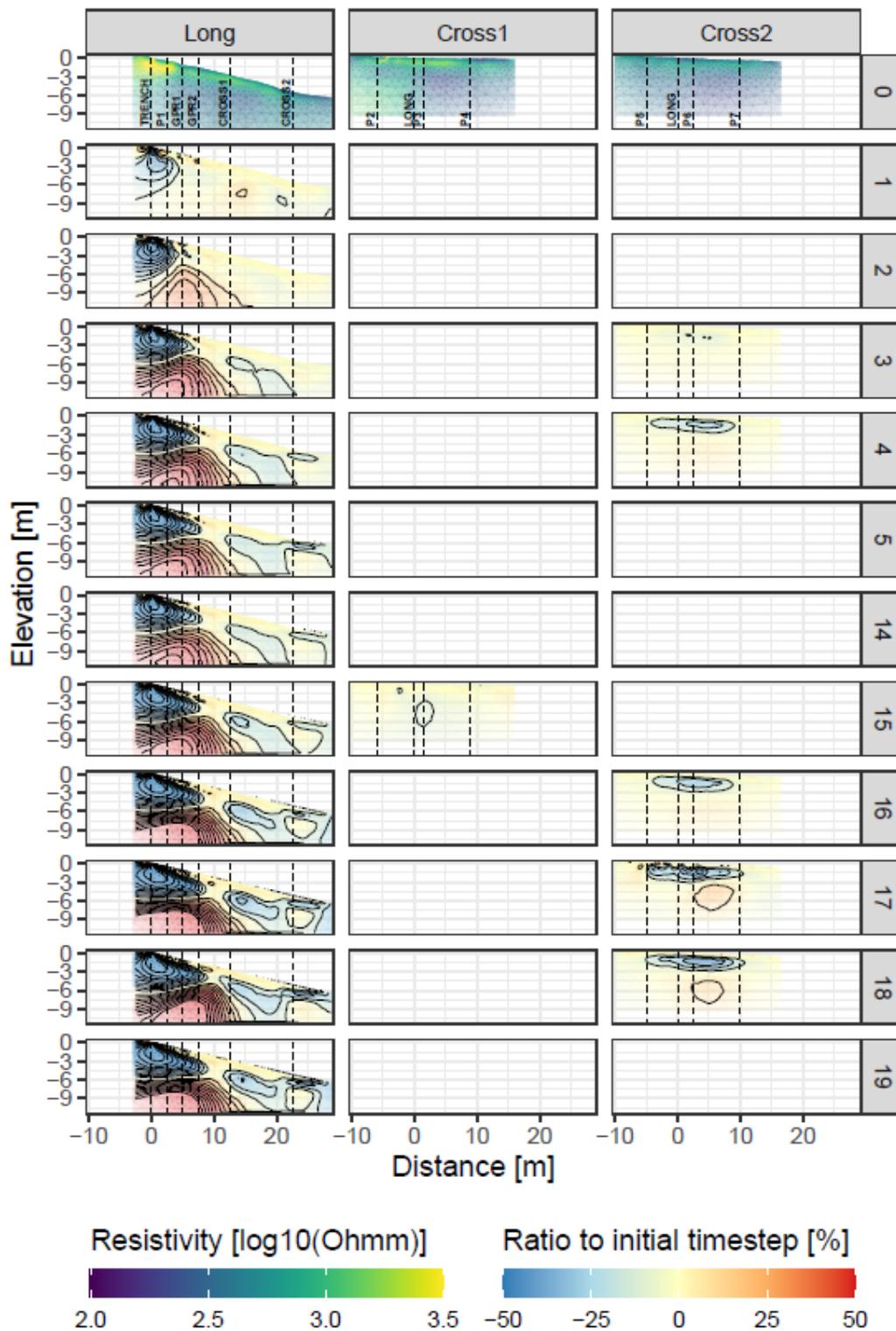


Figure 14: Reference ERT inversion models (snapshot 0) and ratios to snapshot 0 for all other snapshots. Values correspond to resistivity (inverse of EC). Colours are capped at +/- 50%, contour lines are set every 10%.

Note that response of the piezometers and ERT are in accordance in mid-slope. The time-lapse GPR data revealed response velocities of  $0.6 \cdot 10^{-3}$  -  $2.0 \cdot 10^{-3}$  m/s in good agreement with those of the piezometers (of  $1.1 \times 10^{-3}$  to  $2.0 \times 10^{-3}$  m/s), the aforementioned irrigation experiment of phase I (Angermann et al., 2017) and the maxima of Ks (see Fig. 4). We thus state that time lapse geophysics,

particularly GPR in combination with soil moisture observations is well suited to image and quantify vertical and lateral subsurface flow processes.

2.2.1.2.3 Identification of preferential flow occurrence through soil moisture monitoring at 135 soil profiles

In line with the findings reported in the previous section, an event-based analysis of in total 135 soil moisture profiles in phase II project D revealed frequent evidence of preferential flow. Soil moisture profiles were checked for the sequence of their response over depth: sequential response from the surface or non-sequential response (Graham and Lin, 2011). The latter is regarded as fingerprint of vertical preferential flow bypassing top soil sensors, assuming that subsurface lateral flow can be neglected. This analysis revealed a fraction of up to 25% non-sequential events at forested sites, which clearly declined with increasing antecedent soil moisture content and a seasonal trend in NSR occurrence with summer and autumn showing a higher non-uniform flow occurrence (NSR) than in winter and spring (Fig. 15). Phase II Project D also corroborated the feasibility of a Stokes flow infiltration model for simulating sequential infiltration (Demand et al., 2019) and used this to explore process homogeneity landuse unites as function, defined based on their and geological setting (compare also 2.2.2.3).

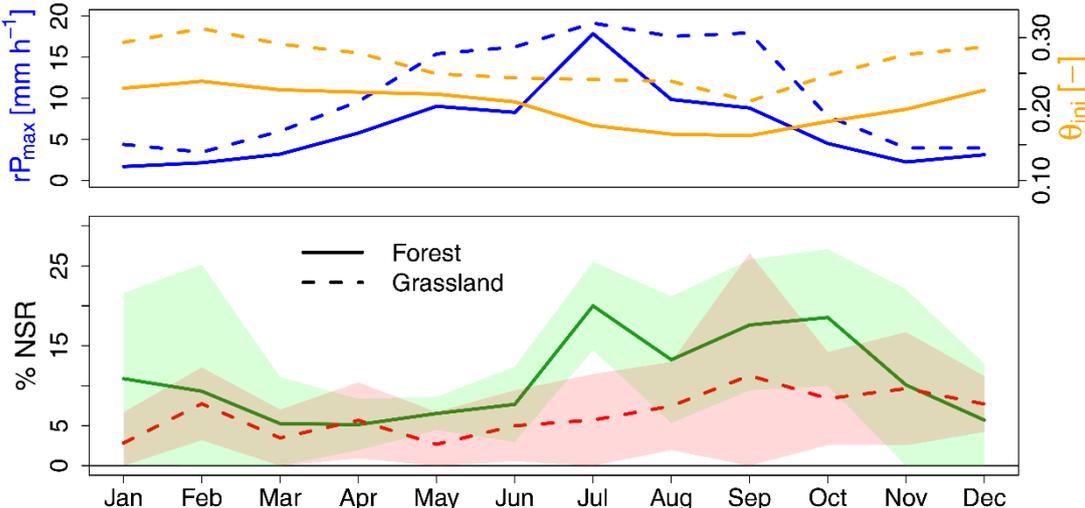


Figure 15: Monthly mean “maximum rain intensity per event ( $rP_{max}$ )”, initial soil moisture content  $\theta_{ini}$  (upper diagram) and fraction of NSR events for the two land covers (lower diagram). The solid lines represent the forest and the dotted lines the grassland response. The shaded areas in the lower diagram show the standard deviation between the single years for each month. Source: Demand et al. (2019).

High wetting front velocities ( $v_{max}$ ), the second indication of PF in soils, were found for Marl profiles, especially when reaching depths of 30 to 50 cm (Fig 16). Lowest wetting front velocities were observed for Sandstone forest. Fast wetting front velocities were found to be independent of initial water content. A reason for the high wetting front velocities in the Marl can be seen in the structure formation in these clay-rich soil horizons (30-50 cm) which increase macroporosity. This demonstrates the danger of treating especially clay soils in the vadose zone as a low-conductive substrate, as the development of soil structure can dominate over the matrix property of the texture alone. Compared to 1D capillary flow velocity, the observed wetting front velocity was faster during 73-89% of the infiltration events, depending on landscape unit and shows that infiltration cannot be evaluated by properties of the soil

matrix alone. The wetting front velocity also showed a seasonal pattern, which is more pronounced in the grassland than in the forest soils. This could indicate that grasslands show more vertical flow, potentially created by the soil fauna.

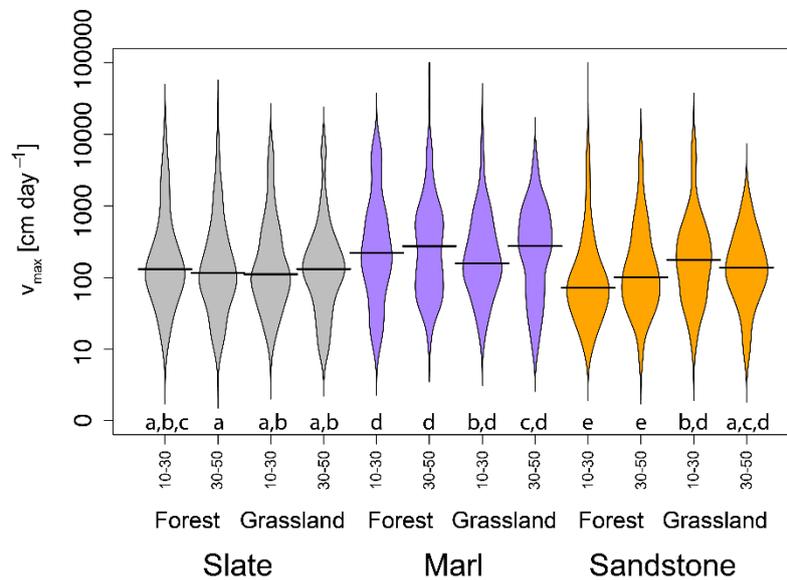


Figure 16: Violin plot of observed wetting front velocity ( $v_{max}$ ) for the six landscape units (colours) and two depths (10–30, 30–50 cm). Same letters below the plots symbolize no significant difference ( $p < 0.025$ , Dunn test, two-sided, Benjamini–Hochberg correction). Source: Demand et al. (2019.)

### 2.2.1.3 Characterizing water storage, mixing and stream flow release

Particularly within the second phase project D, in cooperation with project G and F, aimed on a thorough understanding of how subsurface water storage and mixing controls magnitude and age of stream flow generation processes. To this end, we combined the analysis of soil data from the cluster sites, with remote products, multiple tracers as well as rainfall runoff data sets as detailed in the following sections 2.2.1.3.1 to 2.2.1.3.4.

#### 2.2.1.3.1 Soil moisture dynamics across scales and landscape settings

The 45 sensor cluster sites with 135 soil moisture depth profiles, and around 3-5 years of data, allowed classification of soil moisture responses in high spatial and temporal resolution over time scales longer than common experiments. Distributed soil moisture time series showed a strong rank stability as well as distinct differences in total soil moisture storage, spatial variability as well as dynamics between the different geological settings (Fig. 17). Spatial variability of near surface soil moisture in the Schist geology was in forested landuse generally rather stable, while a strong increase with median wetness moisture was observed in grassland (Fig. 18). This systematic was however not that clear in the other geological settings.

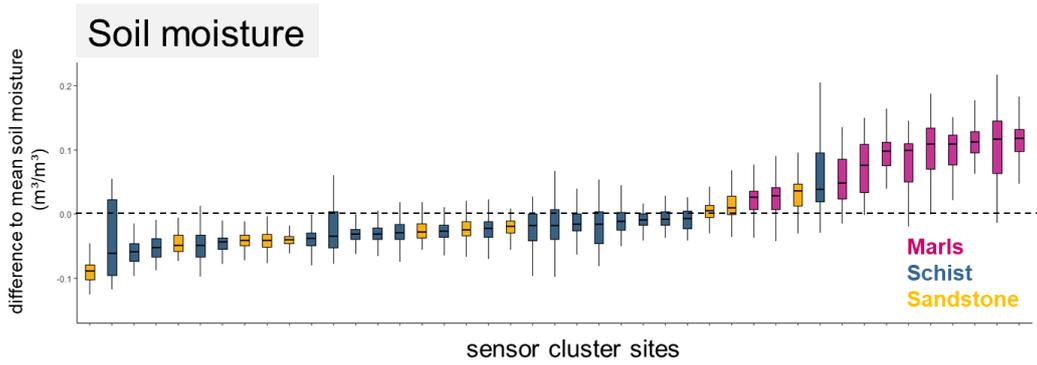


Figure 17: Time rank stability plot of soil moisture at the 45 sensor cluster sites (to be published by Project G).

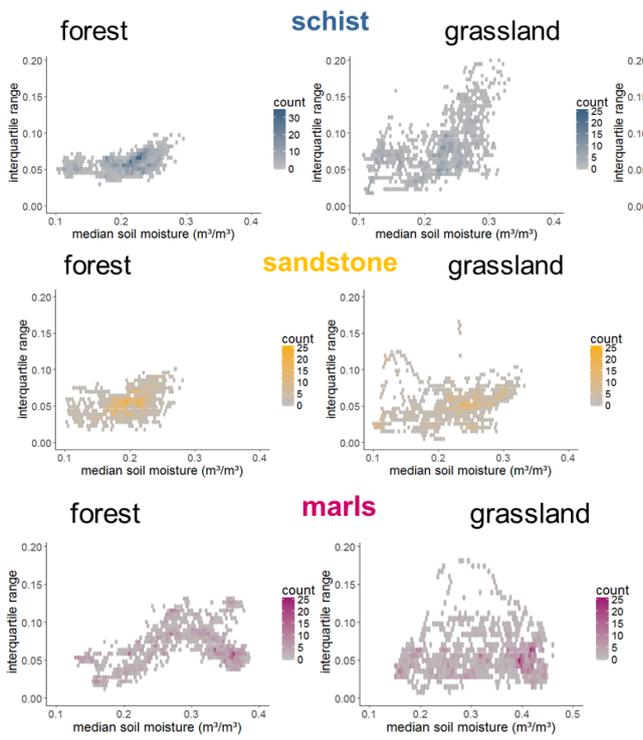


Figure 18: Differences in the relationship of median soil moisture and its spatial variability (here as interquartile range) with respect to geologies and land use (results to be published by Project G).

Time series of spatial rank correlations revealed furthermore a strong rank stability during winter particularly in the Schist and the Sandstone geologies. Time series changed strongly and reorganised as soils dried out due to evapotranspiration in summer (Fig. 19). The marls region shows a very different behaviour compared to the schist and sandstone regions, as in the marls pattern stability is neither achieved for soil moisture nor for sapflow. We also found that soil moisture patterns are much more stable in forests than in grassland (compare report G).

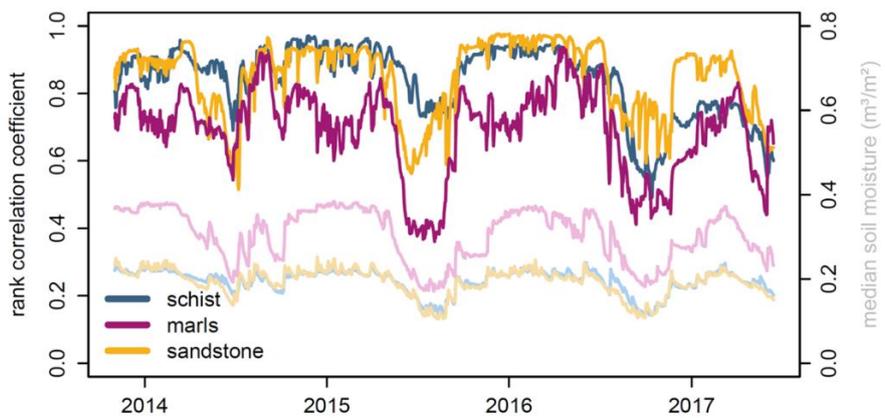


Figure 19: Time series of the rank correlation coefficient between median soil moisture state and the soil moisture state at each respective day across all sites. The corresponding median soil moisture time series per geology is shown in lighter colours. (Results to be published by Project G).

An analysis of soil moisture data in the Devonian slate revealed furthermore that time dependent spatial dispersion of paired monthly averaged soil moisture strongly increased with separating distance. This indicates a distinct spatial dependence/variogram structure of soil moisture time series, which varies seasonally (Figure 20 panel d-f)). Additional clustering of these dispersion functions in time revealed seasonally persistent spatial patterns, and their recurrence in time (Figure 20 a-c, Mälicke et al.; 2020). We propose that these patterns are fingerprints of spatial organised variability, which changes with time due to concurring spatial controls of either the rainfall forcing or transpiration and root water uptake.

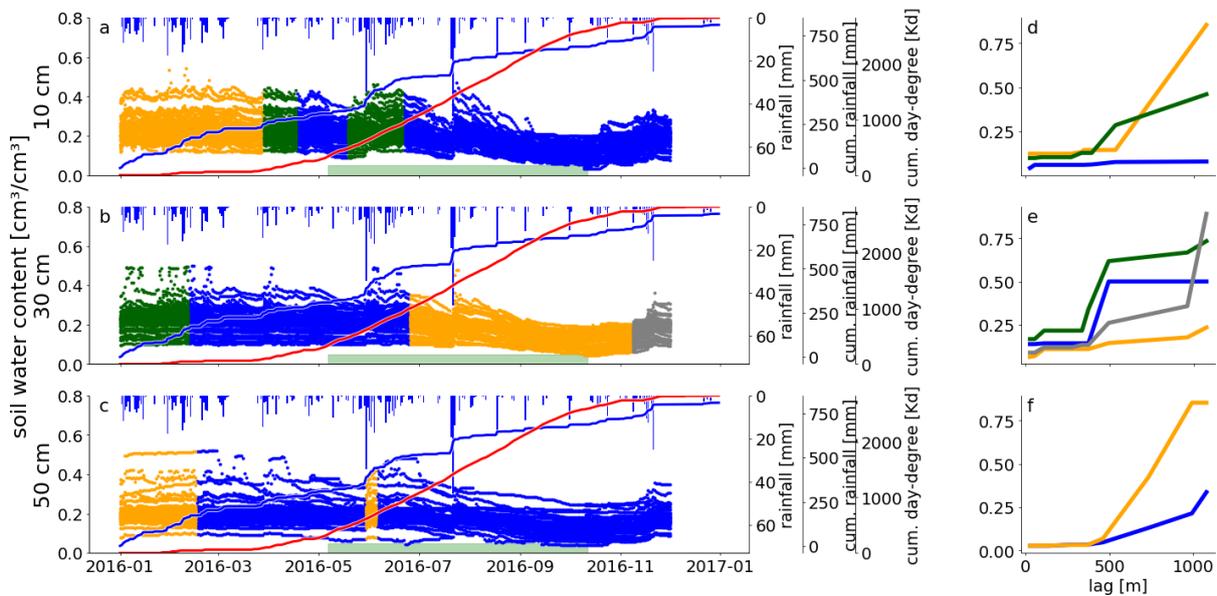


Figure 20: Soil moisture time-series in 2016 in all three depths (panels a–c)), colours indicate the memberships to cluster of similar dispersion. The corresponding representative dispersion functions are given in panels d – f. The red and blue curves in panels a–c represent cumulated degree-day temperature and rainfall respectively. Source: Mälicke et al., (2020).

Complementary to this view, Zehe et al. (2019) developed a thermodynamic perspective on soil water dynamics within phase II project F. They used free energy of soil water to characterize joint gravitational and capillary controls on soil water dynamics. This allowed derivation of a new system characteristic named energy state function that determines the possible range of possible free energy at a given location based on the soil water properties and the height distance to the next stream (compare section 2.2.2.5). The storage equilibrium at zero free energy therein separates storage states of negative free energies corresponding to a storage deficit and dominance of capillarity, from positive ones corresponding to a storage excess and dominance of potential energy. A visualization of soil moisture time-series in their energy state functions revealed distinct differences in the Colpach (slate) and the Wollefsbach (marl) (Fig. 21). Nearly all of the 41 sites in Colpach operated the entire year in the P-regime of storage excess, reflecting the dominance of potential energy and thus topographical controls on soil water dynamics. The 20 sites in the Wollefsbach dropped into a strong storage deficit during the summer period. Note that such differences cannot be inferred from comparing the storage values in both geologies as further detailed in the corresponding final report of phase II project F.

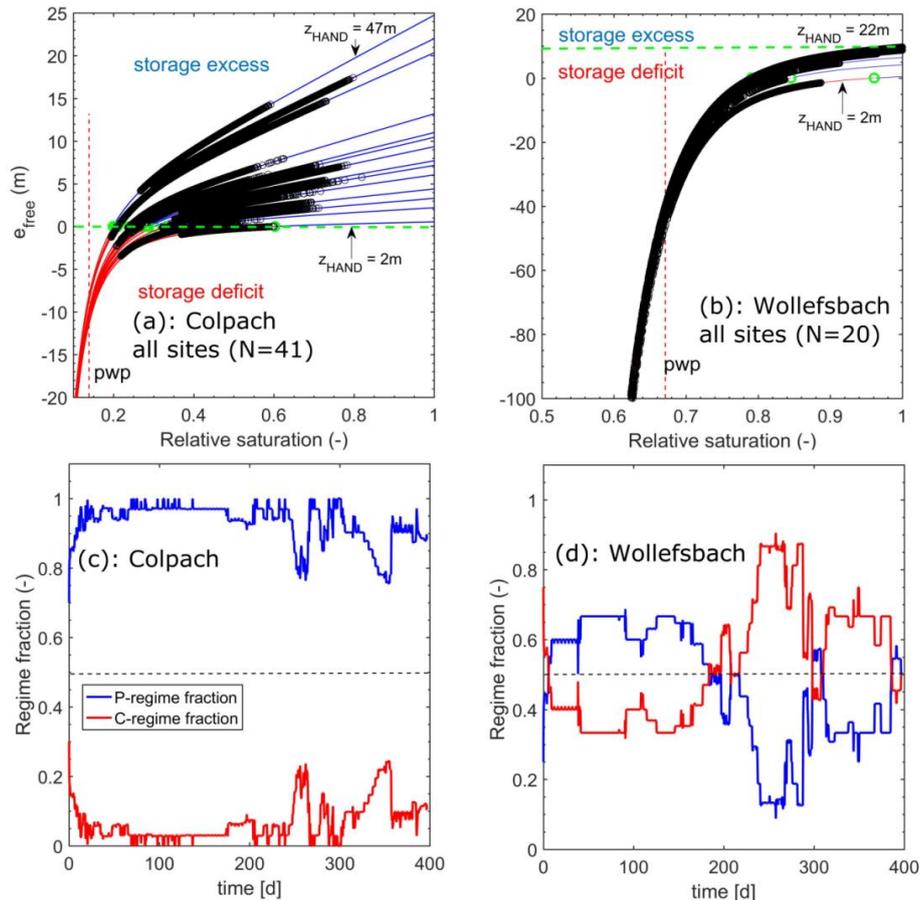


Figure 21: Free energy of all observations in the Colpach (a) and Wollefsbach (b) plotted in their corresponding energy state function (note the different scales). The black circles mark the observations. The horizontal green lines mark the equilibrium of zero free energy. Panel c and d show which fraction of the data set was in the P or in the C regime as function of time. Source: Zehe et al. (2019)

At the basin scale Phase II project D developed a protocol for systematically converting remotely sensed surface soil moisture into physically based maps of root zone soil moisture using Sentinel -1. The corresponding algorithm to estimate soil moisture is based on a multi-temporal maximum likelihood (ML) approach that inverts a direct model of backscattering from bare soils (final report phase II project D). The results show that the combination of remote sensing and data assimilation into a physically based hydrological model allows monitoring changes of water storage across various spatial and temporal scales (Cenci et al., 2017).

#### 2.2.1.3.2 Travel/ transit time distribution of runoff components across scales

Phase I project H conducted distributed sampling of stable water isotopes within soil profiles covering the soil catena to shed light on the interplay of plot scale soil water flow, storage and release. A comparison of the isotope profiles (Fig. 22) revealed considerable heterogeneity of the subsurface processes at each sample site based on the differences between the profiles sampled at each of the snapshot campaigns (red and yellow points). To better understand this interplay, Sprenger et al. (2016) analyzed these data with an inverse modelling exercise using HYDRUS 1d. The simulations of the pore water isotope transport match for most of the d2H profiles the pattern of the field observations. However,

several sites, show for one of the profiles an offset in the simulations to either more enriched values (M\_A, Sa\_A, Sa\_B, Sa\_D, Sa\_E, Sa\_G, S\_S) or more depleted values (M\_E, Sa\_J, S\_G, S\_H, S\_M).

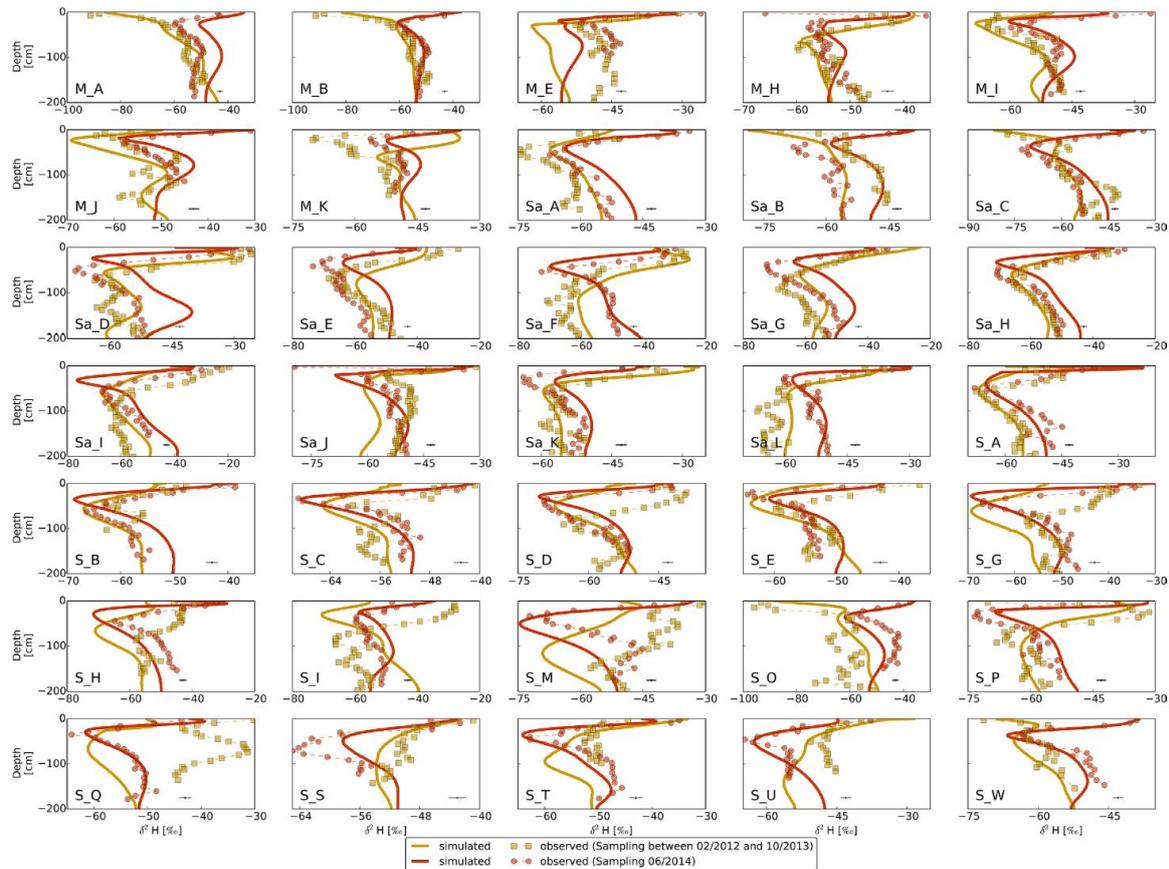


Figure 22: Observed and simulated d2H profiles at each study site. Observed data were used for inverse modeling and the simulated data represent the best fit according to the multi-objective function. Source: Sprenger et al. (2016).

At a few sites also the shape of one of the d2H profiles is not well reflected (M\_K, S\_I, S\_Q, S\_T, S\_U). The average model efficiency for the two d2H profiles at each site were generally higher for the Stagnosols ST (sandstone) and the Arenosols AR (marls) with values for the objective function of 0.81 and 0.80 compared to the Cambisols CM (schist) with 0.73. Since there is no relation between model performance regarding the isotope profile simulations and the slope of the sites, lateral flow does not or just slightly shape the pattern of the pore water stable isotope profiles in the top 1.5–2 m. Even though the CM (schist sites) on the hillslopes in our study hold in the B-horizon horizontally aligned slate fragments in a periglacial slope deposit, a systematic shift of the simulated compared to the observed isotope profiles cannot be observed. However, at the sites S\_I, S\_O, and S\_P (sampled in March after or during snowmelt) peaks of isotopically depleted pore waters in depth below 100 cm were not represented in the simulation indicating possible influences of lateral flow. Similar to the soil moisture simulations, also for the isotope profile simulations, the model efficiency for the CM was lower than for the AR and ST. Thus, the use of the Richards and the advection-dispersion equations to simulate isotope transport and mixing seems to apply less appropriate in the structured CM soils than in the more homogenous soils of the AR and ST. Evaporation fractionation of the isotopic signal was found to be limited to the topsoil and pore waters in marls appeared to be better mixed than in sandstone.

The best model setups were used to infer on the mean transit times of a) seepage water through the upper 2m and b) transpiration for the soils in the three geological settings (Fig. 23). The MTT of seepage water were lowest in the Cambisols on schist ranging between 100 to 500 days and peaked in the Stagnosols on marl, landuse differences were found to be of low importance. Variations of MTT of transpiration are mainly due to seasonal variations in evapotranspiration. The MTTD for the different study sites show that the CM (schist) and AR (sandstone) hold similar seasonal patterns with peaks of high recharge probabilities starting at 40–100 days after the precipitation input, while for the ST (marls), the event waters only contribute to recharge after 400 days after the rainfall (Figure 24).

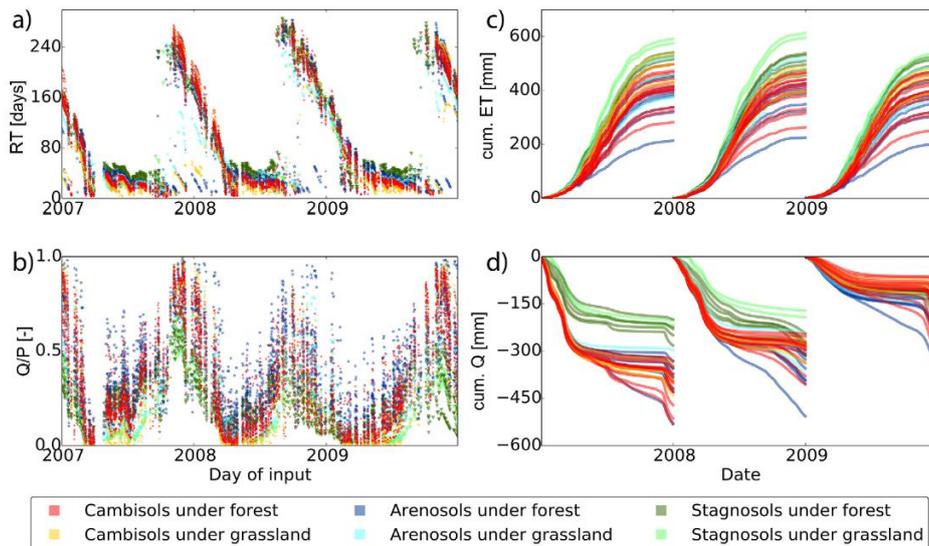


Figure 23: a) The time variant median travel times of the transpired water RT, (b) the time variant recharge ratio Q/P, (c) cumulative evapotranspiration, and (d) cumulative recharge for all 35 study sites colour coded according to the soil types and land use. Source: Sprenger et al. (2016).

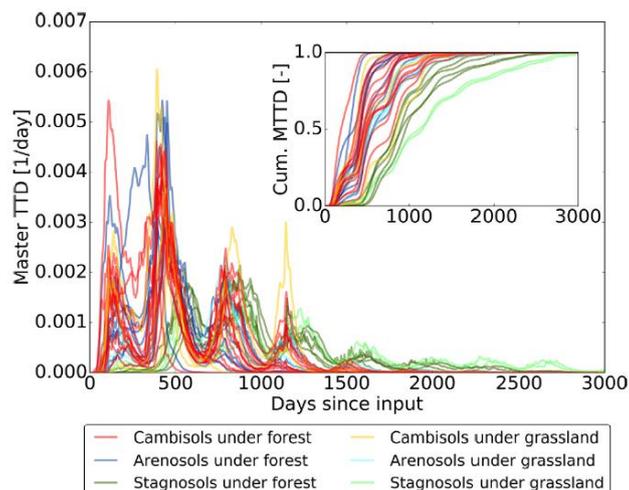


Figure 24: Master travel time distributions (MTTD) for the 35 study sites. Colourcode indicates the soil type and the land use. The inserted graph shows the Cumulative master travel time distributions (Cum. MTTD). Source: Sprenger et al. (2016).

Phase I project H analyzed furthermore travel times of streamflow for the entire Colpach catchment using  $\delta^{18}\text{O}$  concentrations in rainfall and streamflow (compare final report D, group Weiler for the 35

sampling protocol). The available data was analyzed according to the TRANSEP approach (Weiler, 2003) by comparing a two parallel linear reservoirs discharge convolution model to a gamma distribution model within a simulation period of 12 years. This revealed a distinct dependence of the transit time distributions on the length of the data records and of the chosen model, despite of the fact that both model structures yielded a similar fit of the  $\delta^{18}\text{O}$  concentrations in the stream. Mean travel times for the entire Colpach (19.4 km<sup>2</sup>) and the Weiherbach headwater (0.42 km<sup>2</sup>) were 17.8 y and 6.5 y using the gamma model, compared to 49.3 y and 41.1 y for the two parallel linear reservoirs.

Due to these inconsistencies, phase II project D expanded the efforts to assess reliable transit time distributions by combining high frequent, flow weighted sampling of deuterium with a tritium data set collected in the Weiherbach headwater catchment (Fig. 25). To this end, they used an advanced time dependent model concept based on several StoreAgeSelection (SAS) functions for stream flow and evapotranspiration. Aged ranked storage, which is the integral of the transit time distribution over all ages, needs thereby to be inferred from solving the Master equation, i.e., the catchment water balance for each time and each age. This can either be done by using conceptually modelled or observed discharge and evapo-transpiration data. Rodriguez et al. (2021) used a combination of three gamma shaped SAS functions with time dependent weights, which were parametrized either on the deuterium, tritium or on both data sets (Fig. 26). The combined use of both tracers considerably reduced the uncertainty in the inferred time dependent transit time distributions and revealed a median age of  $2.38 \pm 0.15$  y. This is clearly younger than the mean travel times inferred on  $\delta^{18}\text{O}$  and is interestingly in the same order as the mean travel time of seepage water into 2m (compare Fig. 24). We thus advocate a multi-tracer approach to shed light on the age distributions of stream flow. Particularly <sup>3</sup>H contained information that is not in common with deuterium when inferring transport processes to the stream.

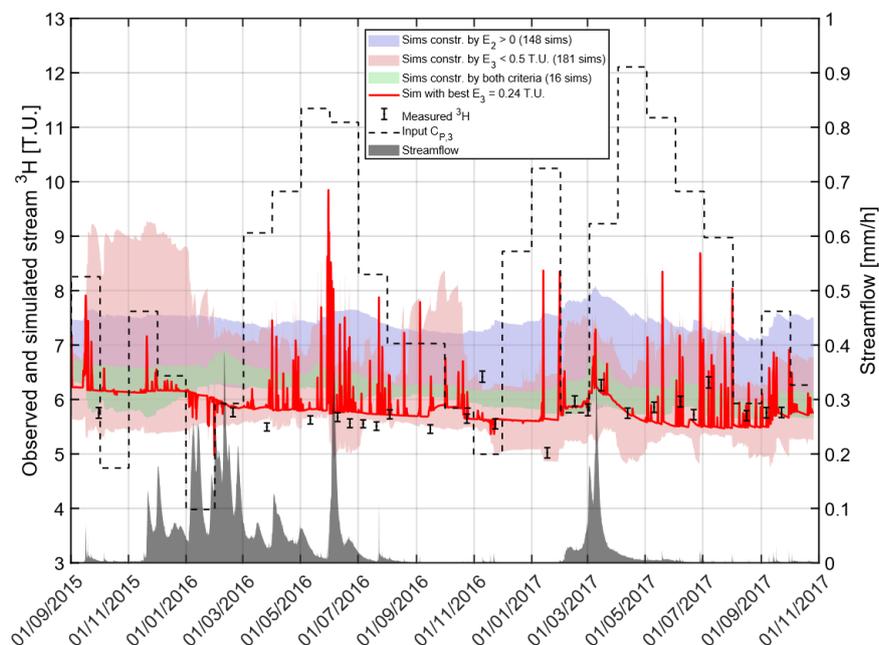


Figure 25: Simulated in tritium concentrations in the stream. E2 is the Nash-Sutcliffe efficiency in deuterium, and E3 is the mean absolute error in tritium units. Source: Rodriguez et al. (2020).

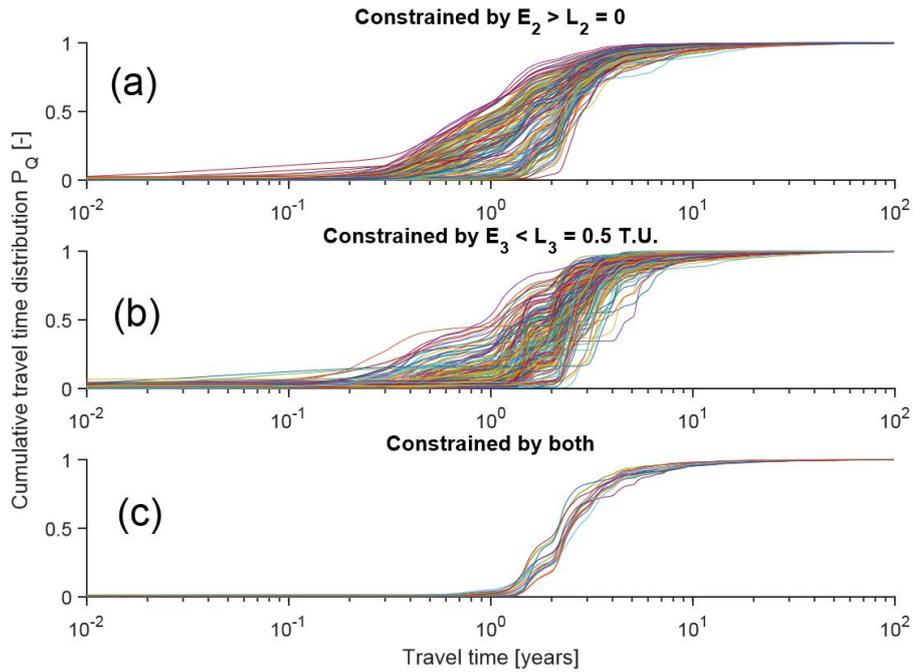


Figure 26: Range in acceptable travel time distributions constrained based on deuterium (a), tritium (b) and both isotopic tracers (c). Source: Rodriguez et al. (2020).

Project D showed furthermore that complementary to the commonly used major elements in hydrological studies, trace elements provide valuable information about flow and transport. This is because these are mineral specific and subject to stronger fractionation during water/rock and soil interaction processes than major elements (Fig. 27).

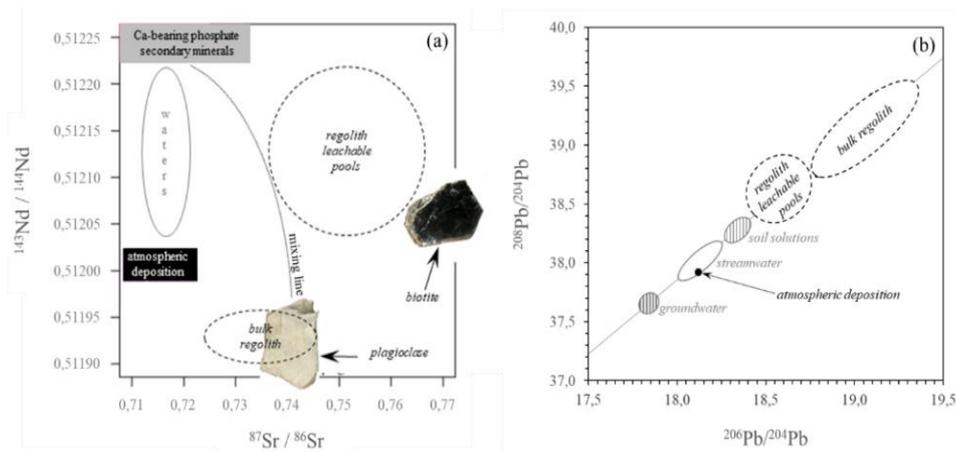


Figure 27: Sr and Nd stable isotope ratios (a), Pb stable isotope ratios of atmospheric deposition, regolith material and water compartments in the regolith of the Colpach catchment in Luxembourg (b). Sr, Nd and Pb isotope ratios in all water types flowing inside the regolith are mainly controlled by the atmospheric deposition and some secondary mineral phases that belong to the preferential flowpaths. Source: Hissler et al. (2015).

Hence, different lithologies and even regolith layers may release typical trace element compositions into solutions. An analysis of Sr-Nd-Pb isotopic compositions for approximately 100 samples (20 water and 76 regolith samples), revealed that any change in isotopic ratios is solely controlled by mixing of different sources. Specifically we found that a large Pb isotopic ratio contrast existed between the groundwater end-member and the subsurface end-member (soil solutions), while streamwater being located in



parafluvial flow paths and stream channel. The spatial distribution of event discharge contributions was primarily controlled by topography, but the signal was modulated and shifted downstream. The discharge patterns were hierarchically controlled, with large-scale signals shaped by topography being superposed by processes at consecutively smaller scales in the riparian zone. Our experiments also showed that stream reach discharge patterns can be used to identify internal catchment processes to a certain extent. The patterns are the fingerprint of processes and structures in the catchment, the riparian zone and the streambed, however, the informational value of processes further away” from the stream channel strongly depends on the degree to which the catchment signal is altered in the riparian zone. Even if under certain conditions the riparian zone and stream channel may be the only factors controlling the discharge patterns, obscuring all information about more distant processes, the relative importance of different processes can always be evaluated. Regarding the methodological approach, we found that the application of differential salt dilution gauging in combination with radon as environmental tracer was suited to identify spatial discharge patterns, but came with a number of limitations, which need to be considered. These challenges include the high variability of end-member background concentrations and degassing rates, which vary strongly with turbulent conditions. Those challenges apply to most applications of radon as groundwater tracer, but are especially pronounced at small scales and for heterogeneous streams. Nevertheless, the distinct concentrations in surface water and groundwater are an advantage especially at small spatial scales and make radon a powerful tracer for surface water-groundwater interactions. In combination with other environmental or artificial tracers to quantify degassing rates, absolute gain, or more clearly define different end-members, radon is a cheap and flexible tool to investigate groundwater discharge into small streams.

Complementary, phase I project A explored the potential for terrestrial diatoms to trace surface connectivity in the hillslope-riparian zone-stream continuum (Klaus et al., 2015). Diatoms (Bacillariophyta) are one of the most common and diverse algal groups (app. 200.000 species,  $\approx 10\text{-}200\ \mu\text{m}$ , unicellular, eukaryotic). In the schist-dominated Weierbach headwater catchment ( $0.47\ \text{km}^2$ ), terrestrial and aerophytic diatom abundance in the stream increased during all 11 sampled events. A higher percentage of terrestrial and aerophytic diatom species was observed in samples collected from the riparian zone and hillslopes than inside the stream (Fig. 29). This flushing of terrestrial and aerophytic diatoms during events suggested a rapid connectivity between the soil surface and the stream. A comparison of diatom transport to two-component hydrograph separations using stream water chemistry and isotope data revealed marginal contributions of hillslope overland flow to stream water, whereas soil riparian water played a major role during events with both wet and dry antecedent conditions.

Motivated by these findings, phase II project G focused on a detailed experimental characterisation of surface and subsurface connectivity. Surface connectivity was investigated using a newly designed system to monitor water levels consisting of a time-lapse camera, a gauging plate and an image analysis tool, which was installed at 71 locations in the entire Attert catchment. The camera offers the use of RGB imagery during daylight and infrared (IR) imagery during the night. The idea to measure water flow velocity based on the bending of a fibre, turned out to be unfeasible. The corresponding dataset was used to analyse space- time dynamics in the river network (Figure 30).

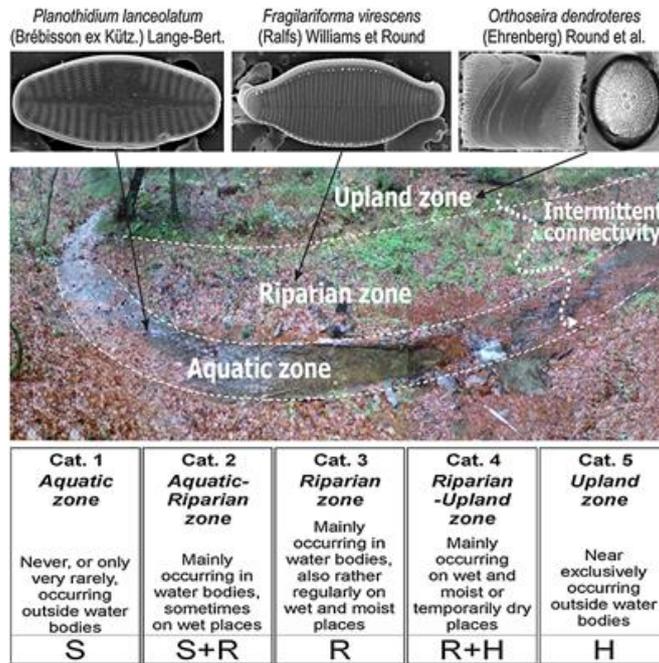


Figure 29: Conceptual classification of diatom occurrence in relation to moisture content (categories 1 to 5) based on Van Dam et al. (1994, cited by Klaus et al., 2015). Related hydrological functional units (aquatic to upland zones), with exemplified diatom species. Source: Klaus et al. (2015).

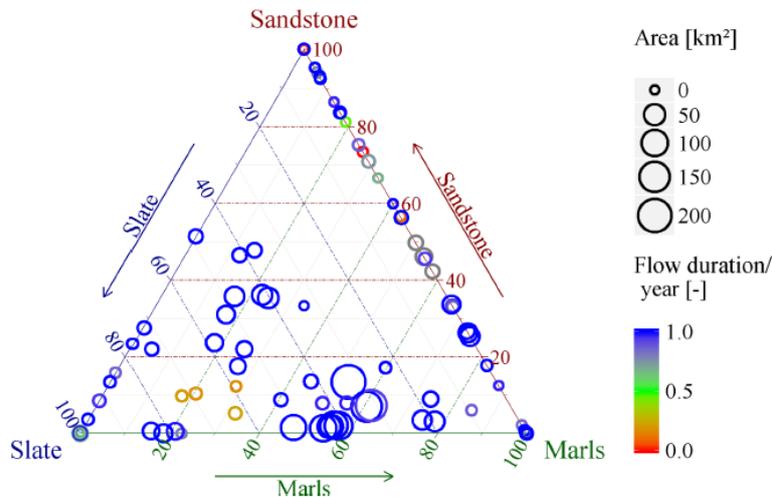


Figure 30: Streamflow intermittency depending on catchment area and geology. Source: Kaplan et al. 2019 (ESSD)

Statistical models were developed to identify the main landscape controls on streamflow intermittency, defined as number of days per year with non-zero flow. Models for the wet periods were mainly driven by the terrain metrics contributing area and profile curvature, which represent a measure for saturation probability. Dry-period models contained relative bedrock permeability and saturated soil hydraulic conductivity as additional predictors, which are a measure of transmissivity and storage capacity of the system in the dry system state. The model for the annual period includes all the predictors from the dry period and additionally track density, which was recognized as a potential indicator of local Hortonian overland flow (Kaplan et al., 2019). As for the dynamic controls the importance of predictors varied strongly between the three geologies (see project report G).

Subsurface hillslope and stream connectivity was furthermore analysed based on the reaction sequence of water levels in the stream and in piezometers in response to rainfall events (Figure 31).

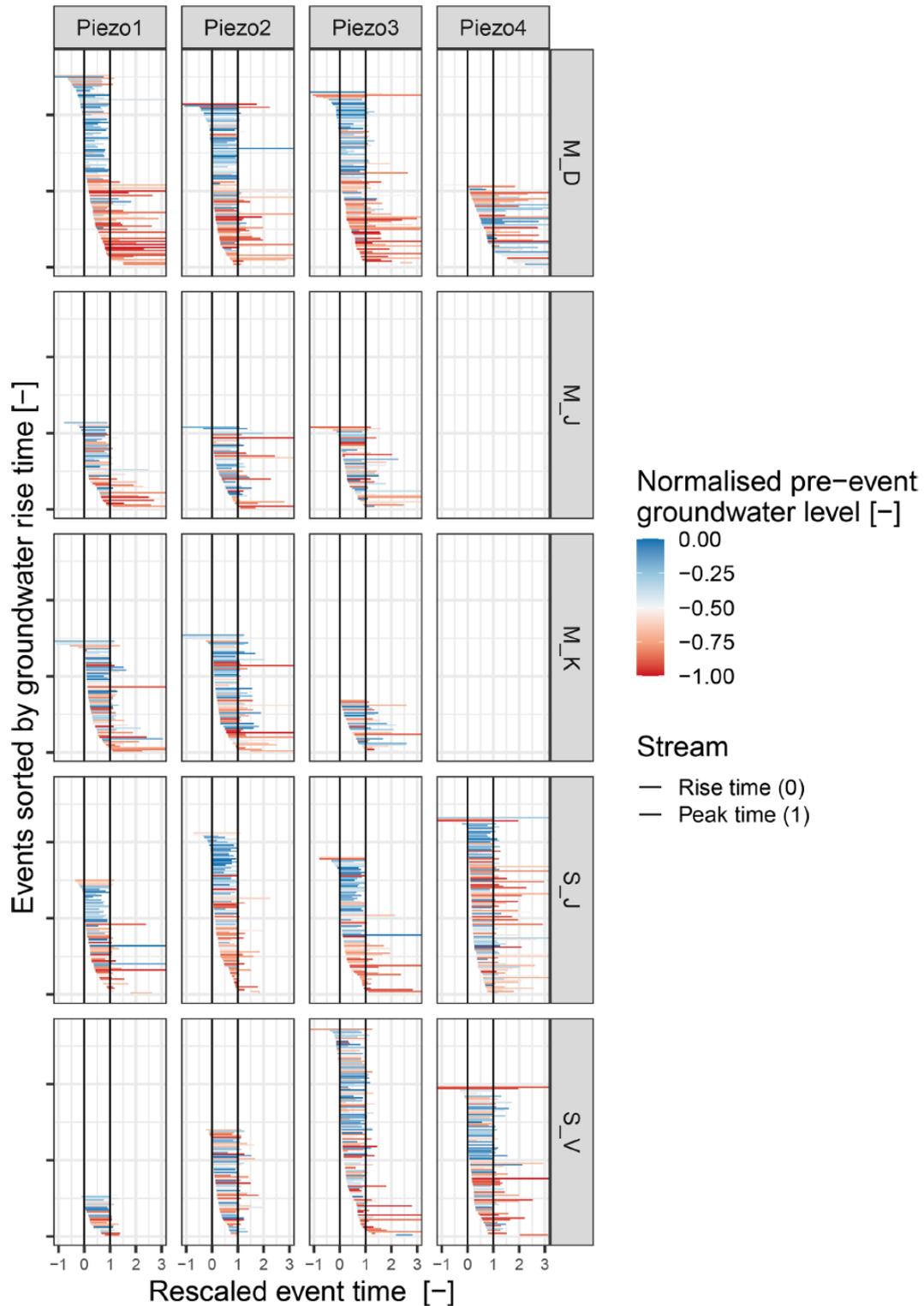


Fig. 31: Timing of groundwater response relative to stream response. M refers to sensor cluster sites in the marls and S refers to sites in the schist region. The two vertical lines at 0 and 1 represent the normalised rise and peak time of each individual streamflow event. Horizontal bars each represent a groundwater event with its individual normalised rise and peak time. Bluish colours indicate high and reddish colours indicate low pre-event groundwater levels. Source: Beiter et al. (2020).

This revealed a close link between shallow groundwater and stream flow dynamics in the Devonian schist geology. More specifically a pre-event GW threshold was found to be helpful to predict the onset or stream flow. This is deemed to reflect filling and spilling of bedrock pools (Beiter et al., 2020), which is consistent with the model results of Loritz et al. (2017) for the Colpach. In contrary to the connection between groundwater dynamics and streamflow generation was found to be rather weak in the marl geology, especially under dry conditions there are strong indications of surface runoff. For more details, please refer to the final report of phase II project G.

In project G phase II we furthermore installed an additional in-stream network of sensors measuring water temperature and electric conductivity with the aim of using this information to infer runoff generation processes and connectivity and investigate distributed streamflow event responses. Temperature and electric conductivity (EC) of the stream water were continuously recorded at 98 sampling sites between July 2015 and June 2018 (Fig. 32). We used Onset HOBO Pendant waterproof temperature and light data logger (Model UA-002-64, Onset Data Logger Manual, 2014) with modified light sensor to measure electric conductivity as proposed by Chapin et al. (2013). Data collection was complemented by regular manual instream control measurements and sensor maintenance. We choose a study design of grouped/nested sub catchments and a high measuring frequency of 10 minutes. Sites were selected with regard to legal and physical access, but with the aim to representatively cover the different stream sizes, catchment sizes, corresponding land uses and lithology classes.

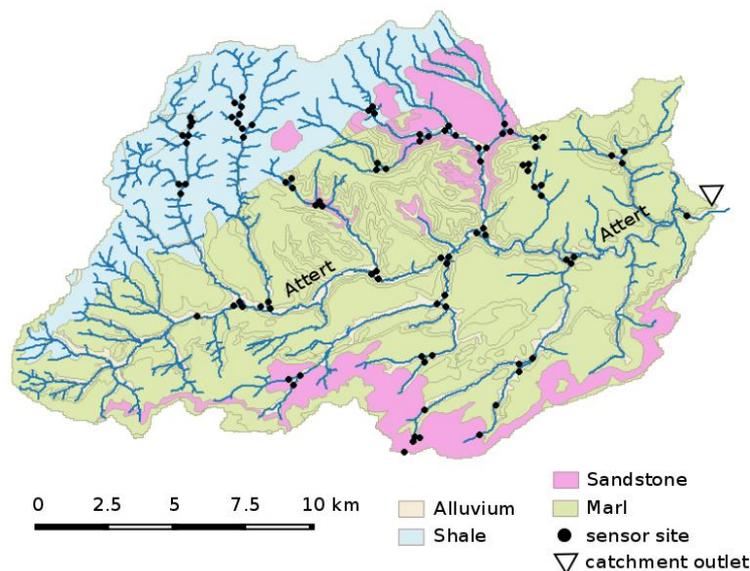


Figure 32: Stream temperature and electric conductivity monitoring network.

The data revealed that long-term patterns of stream temperature and electric conductivity strongly differ among the present lithological units (Figure 33). Stream temperature varied between 0 and 23 °C, following an expected annual pattern with coldest and warmest periods being January and August respectively. We found sandstone sites being cooler than the other groups and showing a smaller diurnal and annual thermal amplitude, while sites within the Marl region south of the Attert stream and sites along the Attert were distinctly warmer, especially in summer. Recorded stream water electric conductivity showed a wide range between 80  $\mu\text{S}/\text{cm}$  and 1800 $\mu\text{S}/\text{cm}$  (Figure 33). Such high values of stream water EC may indicate pollution from agriculture and municipal waste water. Lowest EC values

were recorded at shale/schist sites, where we found a range of  $80\mu\text{S}/\text{cm}$  to  $600\mu\text{S}/\text{cm}$ , with few higher values. Marl sites and some sites along the main Attert stream had significantly higher EC values. Sites along the Attert and within the Marl region showed strong variability in their recorded EC values as well.

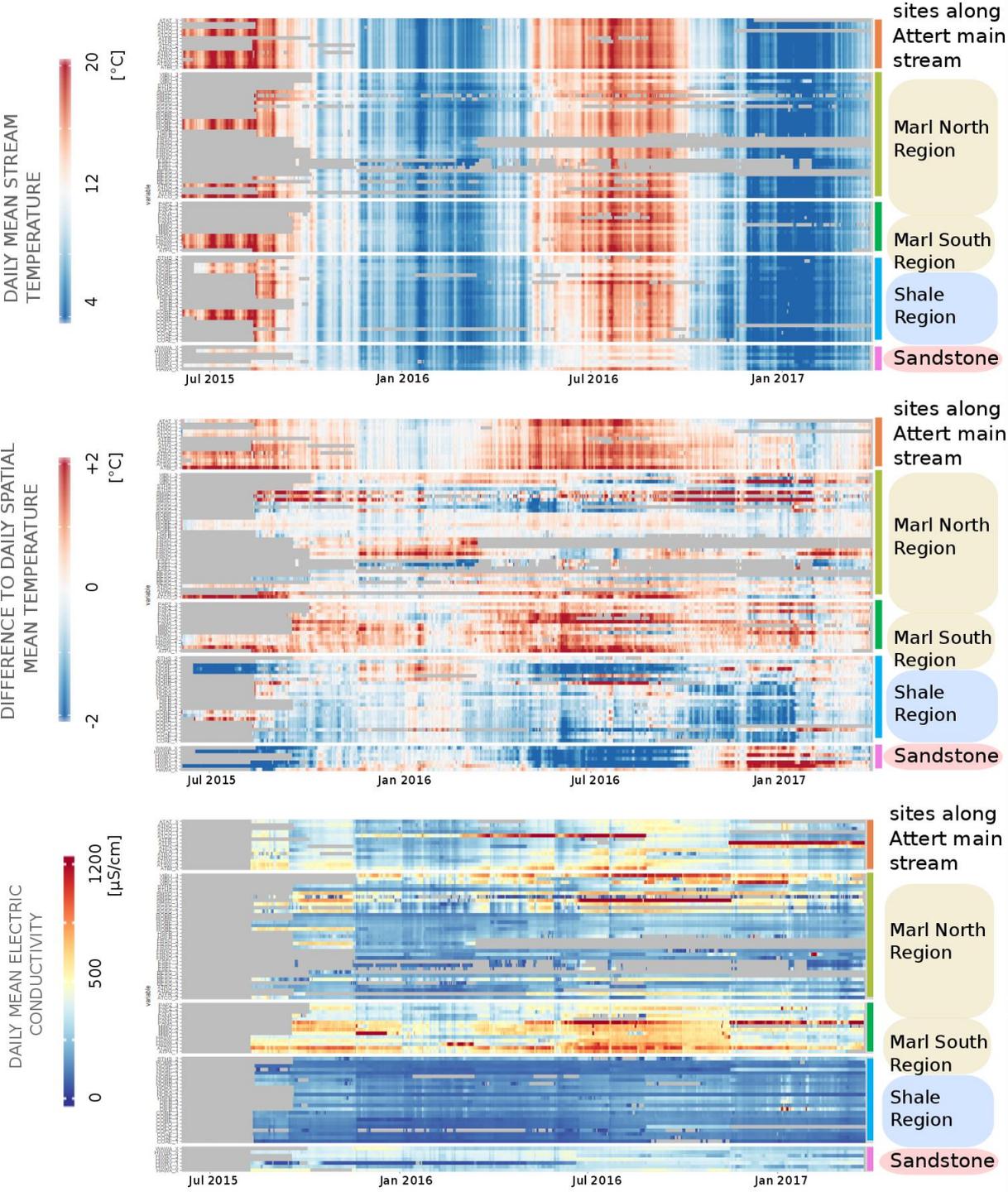


Figure 33: Heatmap of stream temperature and EC time series. Overview of daily mean parameters of the measurement sites. Each row shows the aggregated recorded time series of a measurement site. Sites are grouped by their lithological region. Top: daily mean stream temperature; middle: Temperature difference to daily spatial mean temperature; bottom: daily mean electric conductivity of the stream water. (Results to be published in project G)

In the sandstone catchments, strongly buffered time series of EC and temperature reflect the expected relatively high share of deep groundwater and lower importance of fast flow components for stream discharge. This finding is consistent with the corresponding double mass curves (Fig. 34) and maybe explained by the high permeability and large active storage volume of sandstone aquifer. An opposite behavior is observable in the marl-dominated catchments, where strong effects of seasonality on stream temperature are evident. In addition, there are numerous short-term temperature fluctuations, which correspond to larger rainfall events. Especially the fact that single rainfall events in the summer months are connected to marked increases in stream temperature point towards runoff generation processes on - or close to - the surface. This is consistent with the hydraulic conductivity of the corresponding soils (see also Fig. 4) and the presence of tile drains, making rapid surface runoff of rapid tile drain discharge more likely.

In the shale/schist dominated catchments a moderate seasonal temperature fluctuation is observable, which would point to a groundwater-fed and buffered hydrological system similar to the sandstone catchments. Again, rainfall derived fluctuations are observable in the temperature time series. However, these fluctuations lead to depressions in streamflow temperature at the same points in time, where temperature increases are observable in the marl catchments. An explanation for this could be that runoff generation in the shale-catchments following precipitation events includes larger fast-flow components in depths larger than 50 – 80 cm, where daily temperature fluctuations vanish and water thus equilibrates with these in summer cooler subsoil layers. This runoff generation fits very well with the properties of schist/ shale, which include a high share of rapid flow paths due to macropores and inter-aggregate coarse pores in between the horizontally aligned slates. These subsoil structures would allow water to quickly infiltrate with little belowground retention. Together with the horizontal alignment of the slates this favors a runoff generation where relatively large shares of infiltrating water generates downslope-parallel subsurface stormflow that quickly reaches the streambed.

#### 2.2.1.3.4 Seasonality and storage controls of stream flow generation

Last but not least, we explored the control of water storage on runoff generation at different time scales. At the seasonal scale Jackisch (2015) plotted cumulated normalized discharge release,  $Q$ , against cumulated normalised rainfall supply,  $P$ , at different river gauges (Fig. 34). This revealed a distinctly similar seasonality in marl and schist catchments, a linear increase in the runoff rainfall ratio in the wet season and a switch to very low stream flow release during the vegetated period. The onset and duration of this period of low stream flow release coincides with the onset and ceasing of the vegetation period (Seibert et al., 2017). In contrary to that the sandstone catchment exhibits a linear behaviour the entire year, indicating a large share of deep groundwater in stream flow. Note that these findings are inconsistent with structural catchment the similarities between sandstone and schist areas inferred from GIS and remote sensing data shown in Figure 5. On the event time scale project D found furthermore evidence of a catchment storage threshold controlling the occurrence of the delayed secondary discharge peaks, and different hysteretic relationships between dynamic catchment storage and discharge for single and double peak hydrographs (Martinez-Carreras et al. 2016, Fig. 35. Consistently with that, phase II project showed that the free energy state of soil water stored in the riparian zone

nically explained and predicted the threshold like onset in streamflow generation in the Colpach catchment (schist) and the Wollefsbach catchment (marl), as detailed in section 2.2.2.5.

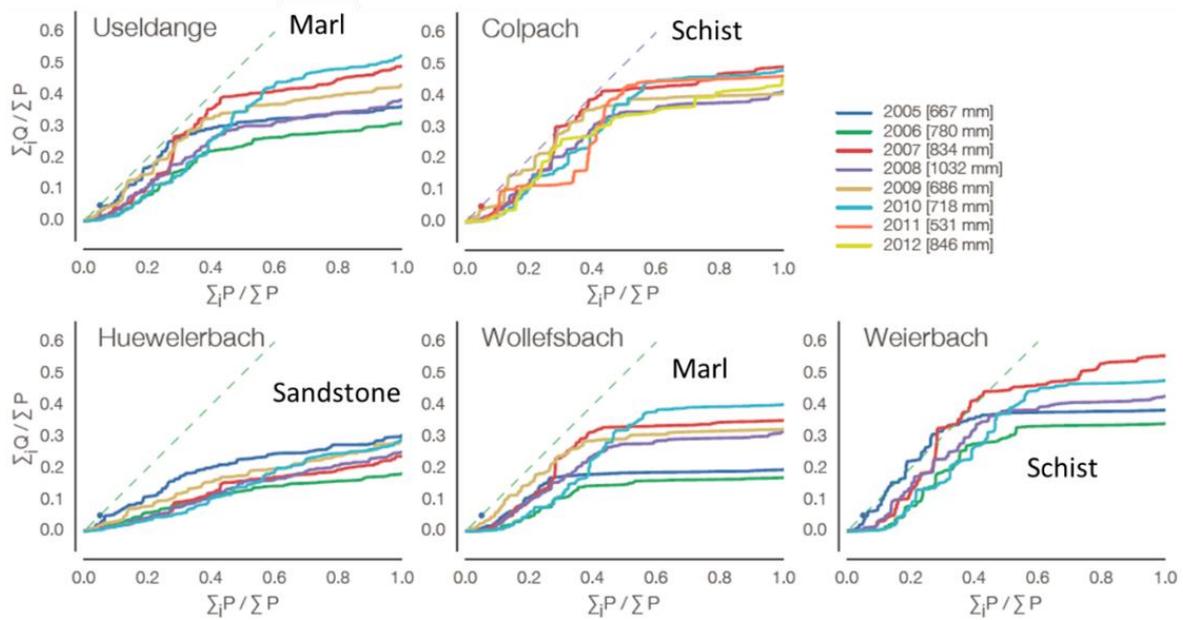


Figure 34: Double mass curves of some of the Attert sub-catchment arranged for inter-annual comparison. As the x-axis is given in relative received precipitation of each year, the timing varies for each line. Source: Jackisch (2015).

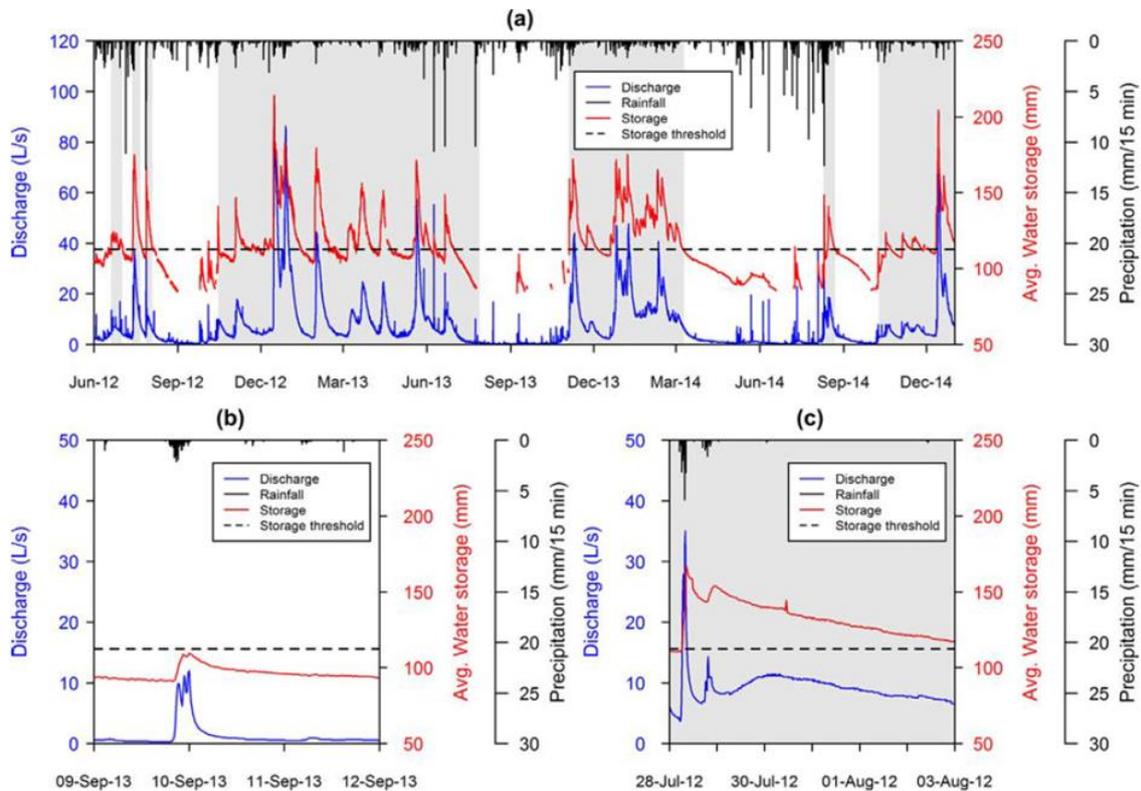


Figure 35: Discharge and average storage for the period June 2012 to January 2014. Shaded zones indicate periods when double peak hydrographs occurred, and the horizontal dashed black line indicates the threshold of catchment storage for double peak hydrographs to occur (i.e. 113 mm), (b) zoom into the 9–12 September 2013 period (single peak hydrograph), and (c) zoom into the 28–31 July 2012 period (double peak hydrograph). Source Martinez-Carreras et al. (2016).

#### 2.2.1.4 Radar based quantitative rainfall estimates

The main objective of phase I project C was to characterize the spatiotemporal variability of rainfall at the catchment scale, using operational C-Band weather radar data from Wideumont and Neuheilenbach. To this end, a sensor network consisting of two Metek MRR vertical radars and six disdrometers was setup to simultaneously observe rainfall intensity and drop size distributions. This was combined with rainfall observations from seven stations operated by the state of Luxembourg and the LIST and the more than 100 rain gauges of the cluster sites (Fig. 36).

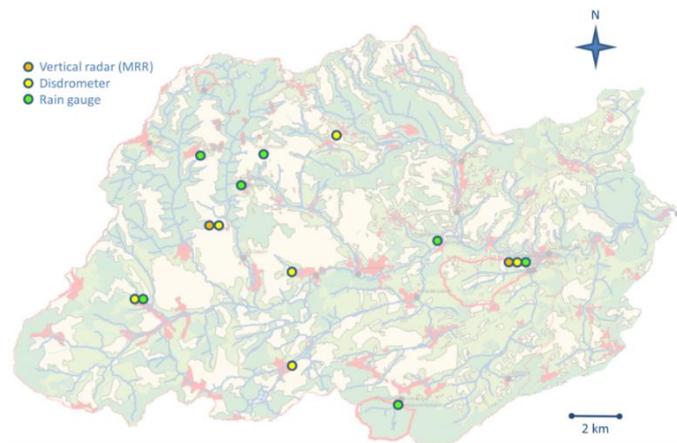


Figure 36: The Attert catchment with sensors related to rainfall estimation. Please note that the more than 100 rain gauges additionally installed at the sensor clusters are not shown.

These data sources were combined along two avenues. Firstly, by means of data-assimilation into the soil-vegetation-atmosphere model system WRF-NOAH-MP (Bauer et al., 2015) using either radar reflectivity and Doppler wind fields or in an upcoming step the new generation of polarization radar systems as further explained in section 2.2.2.4. Secondly, we employed by a geo-statistical merging proposed by Ehret et al. (2013). The basic idea is to condition the radar image to the local rain gauge observations by interpolation the residuals with ordinary kriging, using the variogram of the radar image. During this effort, we also tested the value of disdrometer data to update Z-R relations. Fig. 37 compares spatial rainfall estimates by station interpolation, pure radar rainfall, WRF and the merging method in the verification domain for the passage of frontal rain in the morning of Sep 26. All QPE methods capture the front-like shape of the rainfall field. However, when taking the radar image (upper right panel) as the 'truth', the WRF QPE field appears too small and underestimates rainfall. The merging field preserves the overall structure and extent of the radar rainfall field, but rainfall magnitudes resemble those of the station interpolation field, at least in the vicinity of stations (where the kriging variance is small). In order to assess the quality of the QPE methods in quantitative terms, we computed the mean absolute error between station observations (the 'ground truth') and its related QPE field. For those methods making use of station data (station interpolation and merging), we did so in a leave-one-out cross validation approach. The model-based QPE consistently showed highest errors, followed by pure radar. The merging algorithm performed superior for stratiform and convective rainfall events. In phase II the WRF based QPE was not further developed, while the data based QPE was refined using measures from information theory (Neuper and Ehret, 2019) (see section 2.2.2.6). This effort cumulated in a radar based precipitation climatology, which is now available for the entire Attert basin (Fig. 38). The latter was used

in two studies exploring the feasibility of an adaptive model complexity (Ehret et al. 2020; Loritz et al., 2021) to test our central CAOS hypothesis H3 (compare section 2.2.2).

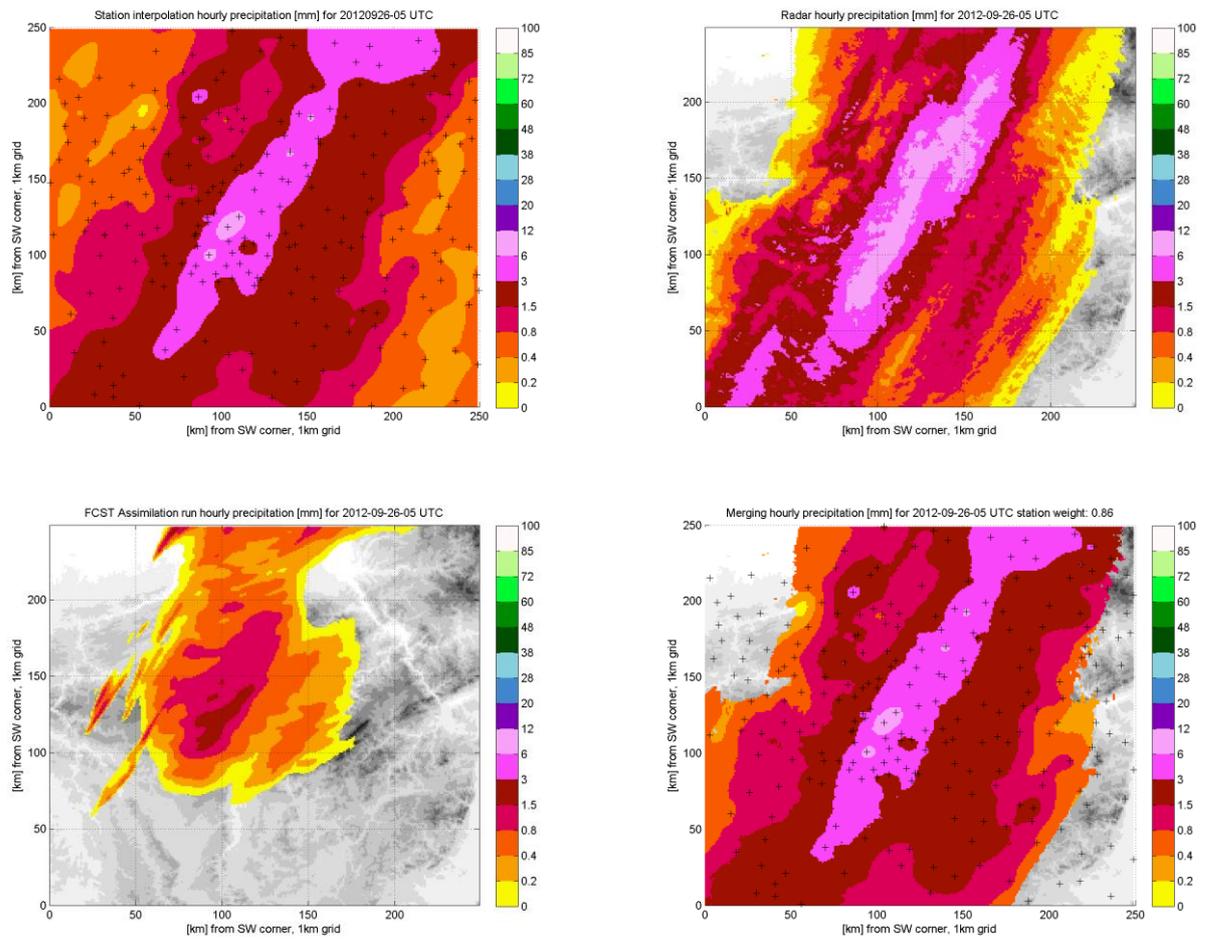


Figure 37: Spatial rainfall estimates for the verification domain (Luxemburg-centered) for Sep. 26, 2012, 04-05:00 UTC. Upper left: Station interpolation, upper right: data from C-band radar Wideumont, lower left: WRF rapid update cycle assimilation run, lower right: merging of station and radar data. Source: Bauer et al. (2015).

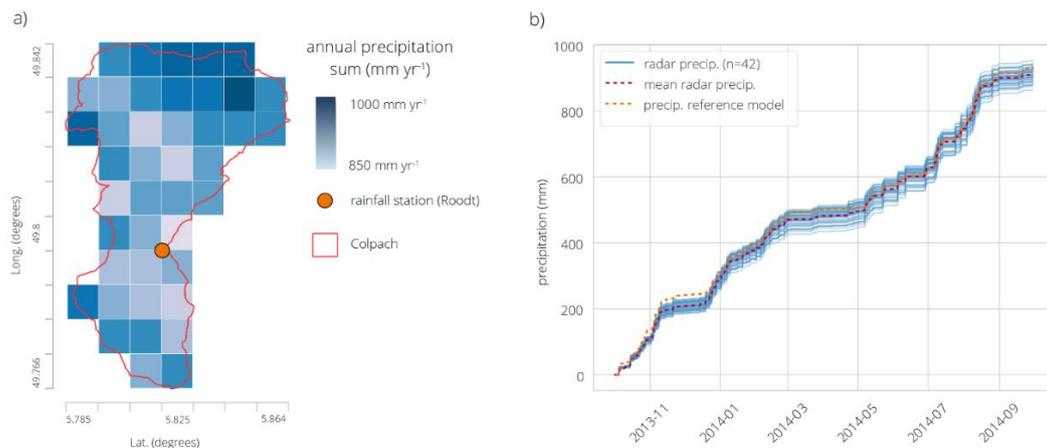


Figure 38: Gridded annual total precipitation exemplarily shown for the Colpach catchment (19.4 km<sup>2</sup>) for the hydrological year 2013/14 (panel a). Time series of cumulated precipitation time series for each grid cell for the hydrological year 2013/14, the corresponding mean of the precipitation field (dashed red line) and the precipitation data from the weather station in Roodt (panel b, dashed orange lines). Source Loritz et al. (2021).

## 2.2.1.5 Characterizing land surface atmosphere energy exchange and feedbacks

### 2.2.1.5.1 Monitoring and experimental campaigns across scales

In our definition land-surface atmosphere exchange and feedbacks include partitioning net radiation into sensible and latent heat and hence evapo-transpiration. We exclude precipitation, although a substantial part of precipitation water stems from regional moisture recycling, particularly during convective events. In phase I, land surface atmosphere energy exchange was experimentally characterized by highly resolved radiation, sap flow, soil moisture and soil temperature data collected at the 46 sensor clusters by phase I project G (compare Fig. 2). Also stable isotope profiles collected at 35 sites by phase I project H provide due to the fractionation information about evaporation (Sprenger et al., 2018, compare Fig. 15). This distributed local information was completed by scintillometer observations along selected valley profiles (Thiem et al., 2017). While it is generally agreed that eddy covariance (EC) observations are the most accurate means of measuring turbulent heat fluxes, these are only representative for a relatively small area around the EC tower. Scintillometers can be used to obtain path-averaged surface fluxes over distances of a few hundreds of meters up to 5 km. Phase I project B thus conducted scintillometer measurements, which were continued at various selected grassland sites in Phase II project C (group Schulz). This was completed by air-borne thermal remote sensing (TIR) yielding spatially highly resolved data on canopy/leaf temperature and soil surface temperature at different time slices as well as satellite based TIR on Modis by phase I project B.

Within phase II these efforts were sustained and expanded along a hierarchy of scales. Phase II project C conducted an extensive field campaign to characterise land-surface energy fluxes and evapotranspiration with multiple methods. The study site was located in Petit-Nobressart (49°46′46.27″N, 5°48′13.07″E, 381.5 m above sea level). Land use type and soil texture were grassland and primarily loam (Fig. 39).



Figure 39: Location and layout of the EC station at Petit-Nobressart, Luxembourg. Source: Brenner et al., (2017).

During a field-campaign in June and July 2015, the site was instrumented with a weather station, an eddy covariance (EC) station and a scintillometer (groups Wulfmeyer and Schulz). Additionally, Brenner et al. (2017; 2018, group Schulz), explored the feasibility of unmanned aerial systems (UAS) for high-

resolution estimation of ET at the (sub-) field scale. An octocopter UAS, equipped with a dual camera setup allowed for the acquisition of optical and thermal imagery with resolutions of 1 cm and 5 cm, respectively. Weather during the field campaign period in July 2015 was characterized by mostly clear sky conditions and high air temperature with daily maxima above 30 °C.

At the scale of the Attert basin site, multiple airborne TIR images were acquired in the phase II project C (LIST) over 22 forest plots during a summer day at 7 July 2016. These images were converted to surface temperature ( $T_s$ ) and to evapotranspiration (ET) maps using the STIC model (Bhattari et al., 2018) together with meteorological data and compared to in situ measurements of transpiration with sap flow.

#### 2.2.1.5.2 Tree transpiration inferred from sap flow velocities: variability and uncertainty

We examined controls on spatial patterns of transpiration in the Attert landscape, inferred from sap velocity and sap flow (Hassler et al., 2018). To identify dominant controls and their possible seasonal dependence, we applied general linear models on daily aggregated sap velocity values, using the following predictors:

- tree-specific: species, diameter at breast height (DBH) and tree height;
- stand-specific: basal area (BA), i.e. stem area related to the ground area;
- site-specific: geology, slope position and aspect.

The spatial patterns were mainly controlled by tree- and site-specific characteristics (Fig. 40). Temporal dynamics of the overall explained variance of the linear models and the relative importance of species was closely linked to the dynamics of potential evaporation, whereas the site-specific influences remained constant over time. This means that the abiotic characteristics of the landscape control transpiration pattern to a certain extent, and this control remains static in time. However, the importance of biotic characteristics, i.e. the landscape-scale patterns of tree species distribution, varies in time and becomes most important during days of high atmospheric demand. Our results suggest that spatial representation of landscape-scale transpiration in distributed hydrological models could be improved by including spatial patterns of tree-, stand- and site-specific characteristics. For spatial sap flow patterns, these influences were considerably larger than the obvious and widely used influences of the potential evaporation and water availability in the soil.

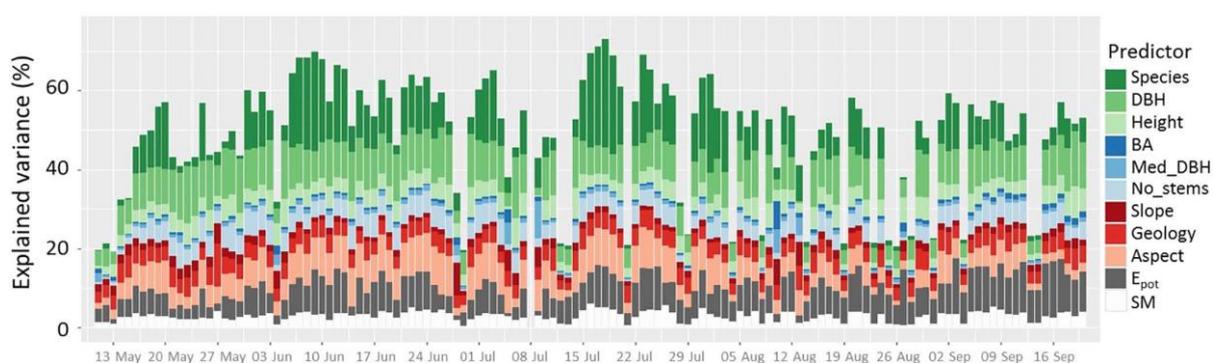


Figure 40: Proportion of variance explained by the different predictors in the daily linear models of spatial sap velocity patterns: 132 daily models from 61 trees at 24 sites.

Consequently, similar to resolving agricultural areas according to crops on a field scale, one could represent the spatial structure in forest transpiration resulting from species and size distributions, but also from patterns due to site characteristics such as geology or topography. This information can be used for model parameterisation or as a part of multi-response evaluation for soil–vegetation–atmosphere transfer and hydrological models. In a first feasibility study, we upscaled available sap flow measurements to stand level transpiration with the help of regression equations relating tree diameter and sap flow using data from stand surveys (Strunck, 2020). Uncertainties in these estimates result from the inherent uncertainty of sap flow at the tree level, from the regression equation and from the neglected proportion of understorey transpiration. Nevertheless, stand-level transpiration can be better compared with other larger scale estimates, for example from remote sensing products or models. We compared stand transpiration from seven cluster sites in the Devonian schist area and nine cluster sites in the sandstone geology (Fig. 41). The results show a larger diversity, which partly related to the site characteristics such as aspect or distance to a stream (and groundwater). Of course, also the forest management at each cluster site affects these differences.

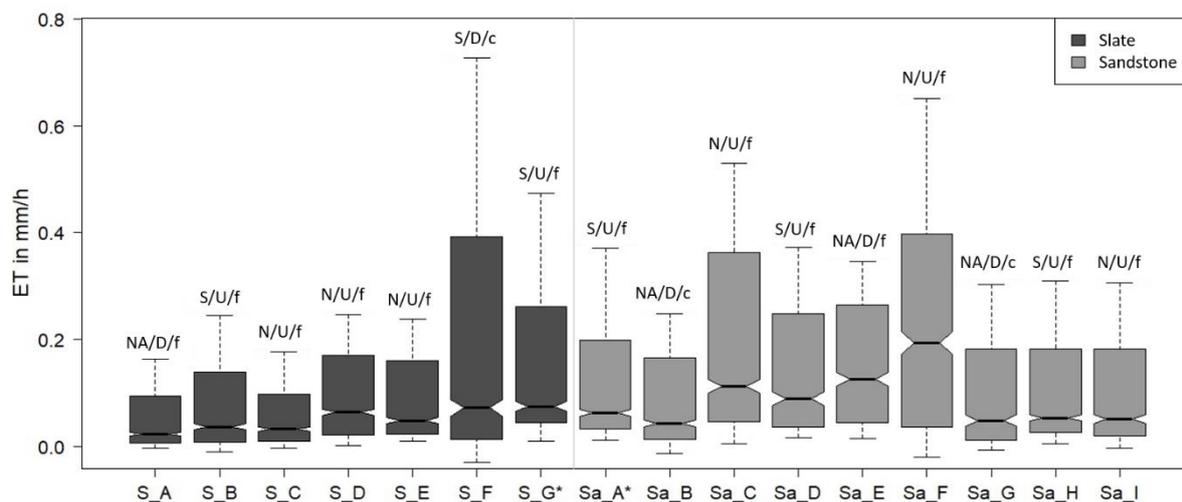


Figure 41: Stand transpiration of all examined cluster sites in June 2015 in the schist/slate and sandstone area. The abbreviated site characteristics in the figure are: Aspect (NA: no aspect, S: south, N: north), slope position (D: downslope, U: upslope), distance to stream/groundwater (f: far, c: close)

In a second step, we compared the stand transpiration estimates to an ET product from EUMETSAT with a temporal resolution of 30 minutes and a spatial resolution of 2000 m x 3000 m. The stand transpiration values represent an area of 20 m x 20 m at the same temporal resolution (Fig. 42). This revealed that the large variability of the individual stand transpiration estimates is hardly reflected in the remote sensing product due its coarser resolution. Additionally, the temporal dynamics differ between the two estimates. As the remotely sensed ET is mainly calculated from meteorological variables (e.g. radiation, air temperature, humidity, wind velocity) and some additional characteristics such as LAI and soil moisture, this product strongly reflects the physical drivers and limitations of ET. Sap flow is the trees' response to these drivers and consequently lags behind the remotely sensed estimate. These differences in spatial and temporal representation must be kept in mind when choosing an ET estimate for a certain purpose.

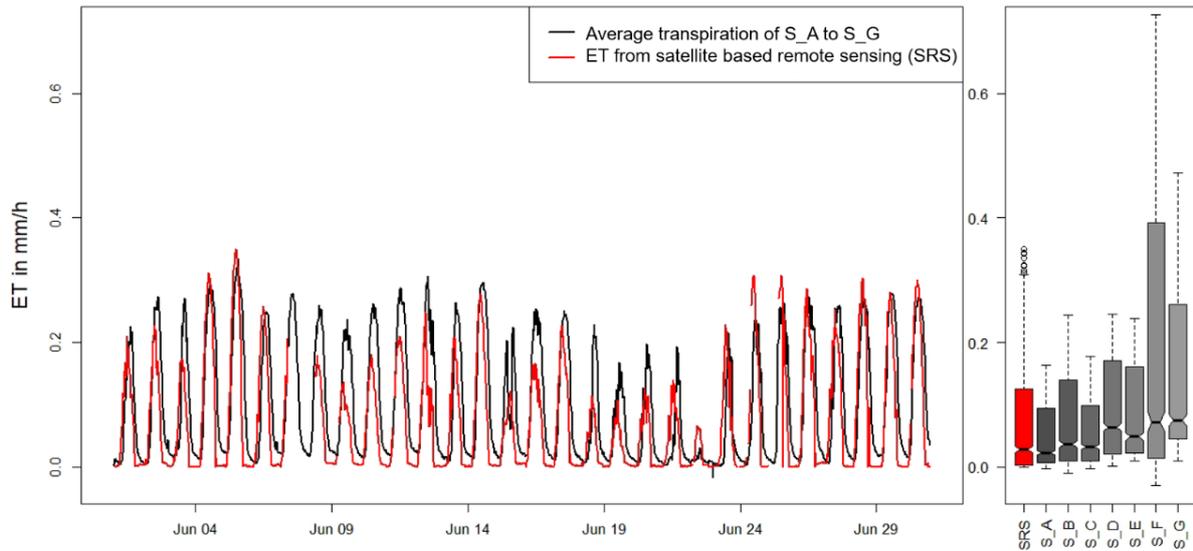


Figure 42. Comparison of (ET) estimates from sap flow measurements in seven slate clusters and from remote sensing for June 2015.

#### 2.2.1.5.1 Inferring root water uptake from soil moisture observations

To further illuminate the interplay of transpiration and soil water dynamics phase II project F took a closer look at the interface of vegetation and soil with different measures. Sap flow (SF) describes the water movement within the tree induced by transpiration at the leaf level, with a potential time lag due to storage capacitance within the stem. From the soil perspective, one expects a diurnal decline in soil moisture, which is supposed to reflect to the root water uptake to supply trees' transpiration. As the translation from root water uptake (RWU), deduced from soil moisture declines, to sap flow and transpiration is not straightforward, we conducted additional highly resolved soil moisture measurements to examine this interdependence. We installed two arrays of TDR tube probes, one each at a cluster site in the slate part of the catchment and at a sandstone cluster site, next to beech trees instrumented with sap flow sensors. The tube probe arrays provide spatially continuous profile measurements in 20-cm depth increments, and thus allow for a detailed analysis of the depth distributions of root water uptake. To this end, Jackisch et al. (2020) developed an automated algorithm to detect and evaluate the step-like soil moisture declines and to calculate the root water uptake. In general, the approach to detect RWU worked well, especially in moderately moist and homogeneously textured soils as indicated by the similar seasonal patterns of sap flow and RWU in Fig. 43. Specifically, we found that during leaf out water is abstracted from deeper soil layers, and that generally the two sites differ with respect to the depths contributing mainly to RWU. A closer look revealed good correlations between RWU and sap flow during spring, while their decline in summer points towards less efficiency of the algorithm to detect RWU during drier periods (Fig. 44). This decline is more pronounced at the pedologically heterogeneous site, which could for example point towards other sources of water contributing to RWU, which are outside our measurement depth, such as cracks in the slate. This suggests that the soil characteristics affect RWU but not SF to the same extent. These findings underpin that SF and RWU are complementary measure of plant transpiration and that the depth of the root store of the same tree species varies among different soil settings.

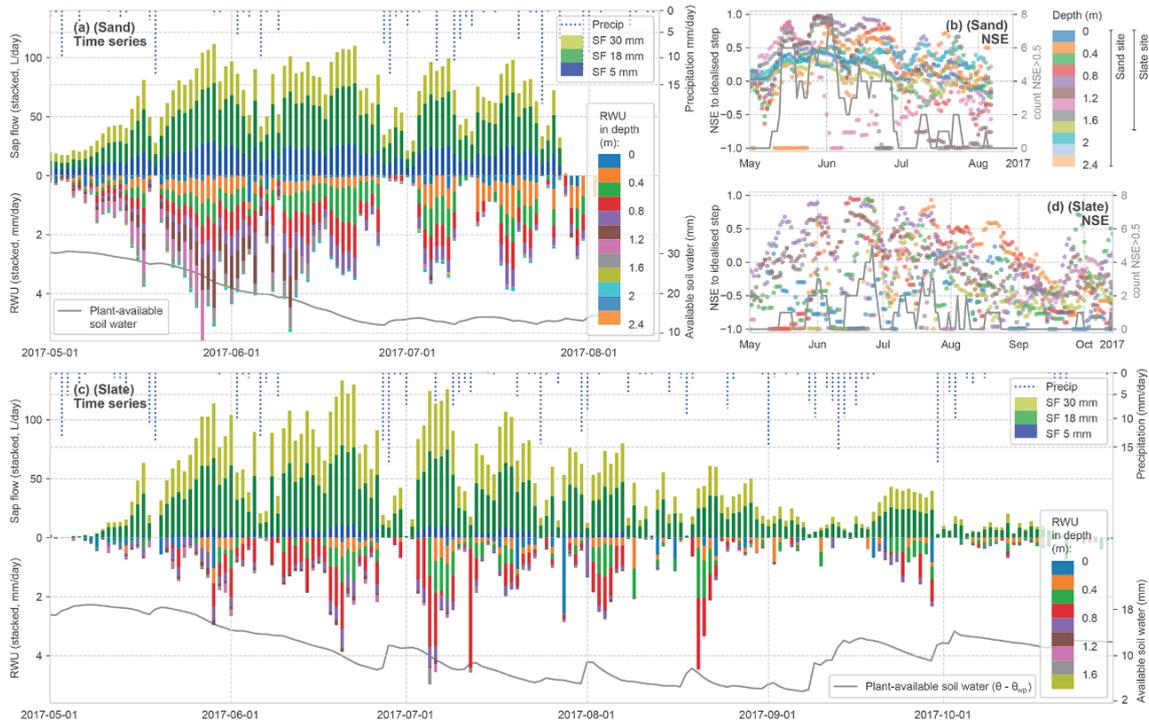


Figure 43: Summary of calculated time series for sap flow (SF) and RWU estimates as stacked daily values, for sandstone and slate. SF (panels (a) and (c) – upper half) is given as volume flux, while RWU (panels (a) and (c) – lower half) is given as the flow of withdrawn water. As an indicator for soil moisture state, we report the plant-available soil water in the soil column as being the difference between the measured water content minus the water content at the permanent wilting point with the grey line at the bottom of panels (a) and (c). In panels (b) and (d), the evaluation of the observed diurnal soil moisture time series to an idealised step is reported for each soil layer at the two sites (rolling 7 d mean of NSE) alongside the daily counts of detected steps with NSE > 0.5 across all layers. High NSE values reflect a high determination of the RWU estimates in stacked bars in panels (a) and (c).

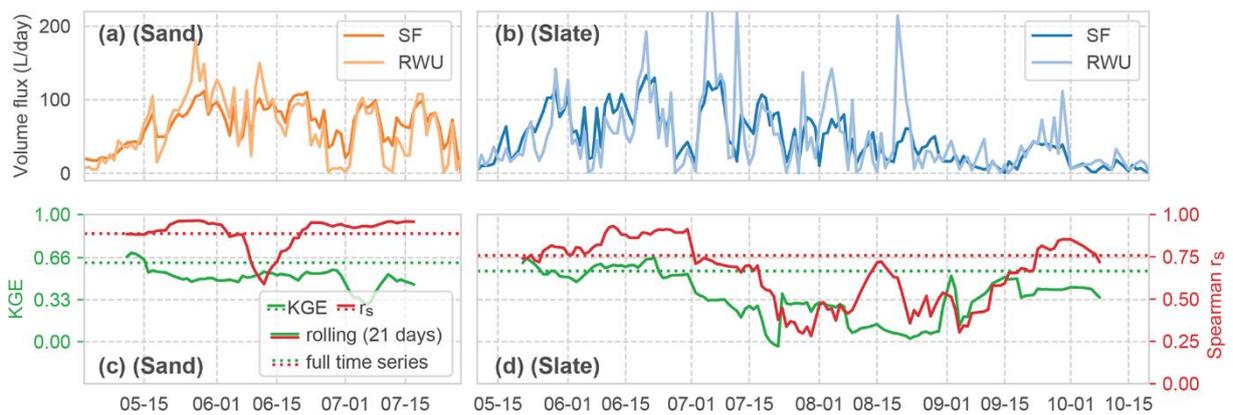


Figure 44: Comparison of the time series of calculated volume fluxes for RWU and SF at both sites (panels (a) and (b) for sand and slate, respectively). Correlations between the RWU and SF time series are shown in panels (c) and (d), both as Kling-Gutpa-Efficiency (KGE) and Spearman's rho. The solid lines for the correlations show a 21-d rolling mean; the dashed lines are the mean correlations for the whole time series.

Another interesting aspect is that the correlations of the two SF time series at the different sites are constantly high over the season (Fig. 45 a), while the coherence among RWU time series is generally lower and more variable (Fig. 45 b).

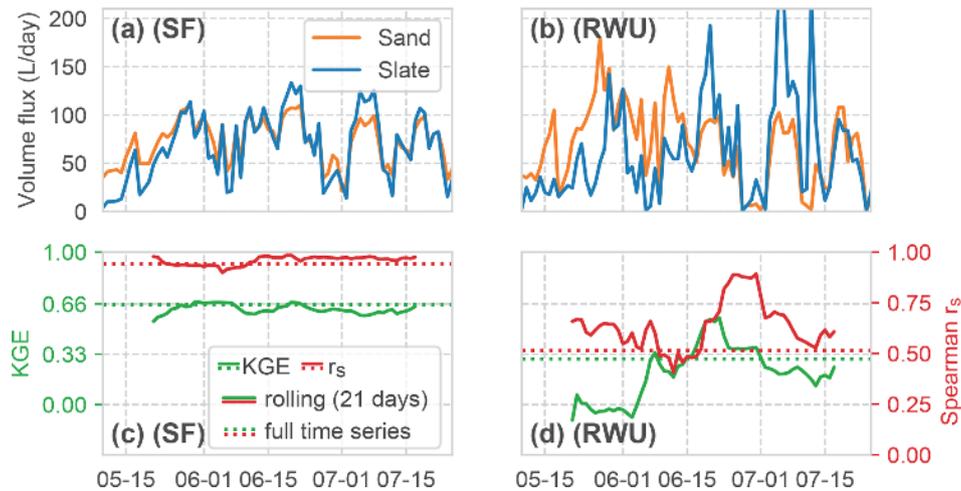


Figure 45: Comparison of the time series of the calculated volume fluxes for RWU and SF between both sites (panels (a) and (b) for SF and RWU, respectively). Correlation of the RWU and SF fluxes between the sites, both as Kling-Gupta-Efficiency (KGE) and Spearman's rho.

Overall, we conclude that profile TDR data in combination with the proposed algorithm are a good means to discriminate RWU from variables depth.

#### 2.2.1.5.2 Detecting sensible and latent heat fluxes using unmanned aerial systems

Information on the magnitude of ET from land surfaces, along with its variability in space and time, is essential for crop water management and other applications in the agricultural sector. However, ET remains elusive to quantify especially when high spatio-temporal detail is desired. A promising alternative is to infer ET based on information on land surface temperature and vegetation properties. This was tested during two field campaigns over temperature grassland sites, using ET models approaches of different complexity ranging from simplistic contextual approaches (making use of the natural variability within the scene) to more physically based one/two source surface energy balance models (OSEB/TSEB). The small spatial extent and the high resolution of UAS-based imagery are superior for model discrimination as compared to satellite remote sensing data. The analyses showed that well-established ET models in combination with UAS-based imagery allow ET estimates agreeing well with EC observations (Fig. 46). In the case of the contextual models, a high spatial resolution of the LST input imagery is, however, decisive for a good model performance. The good performance even of the simpler contextual schemes is especially promising in view of routine near-real-time ET monitoring to provide valuable information on plant water status for e.g. precision agriculture applications.

#### 2.2.1.5.3 Diurnal variation and hysteresis patterns of evapo-transpiration as means of model discrimination

In Phase II project C, Renner et al. (2019) compared the feasibility of different model approaches to simulate the diurnal cycles of evapo-transpiration as well as the hysteresis originating from the phase lags to incoming solar radiation and latent heat flux. The models were constrained with micrometeorological data collected during the summer field campaign and compared against the corresponding eddy covariance (EC) measurements. The study compared four structurally different modelling approaches (Fig. 47): The Priestley and Taylor approach for potential evaporation, for actual evapotranspiration the above mentioned one-source and two-source energy balance schemes (OSEB and TSEB, respectively), which derive the latent heat flux as residual of the surface energy balance and

a parameterization of the sensible heat flux using a resistance description and the surface-to-air temperature gradient as driver.

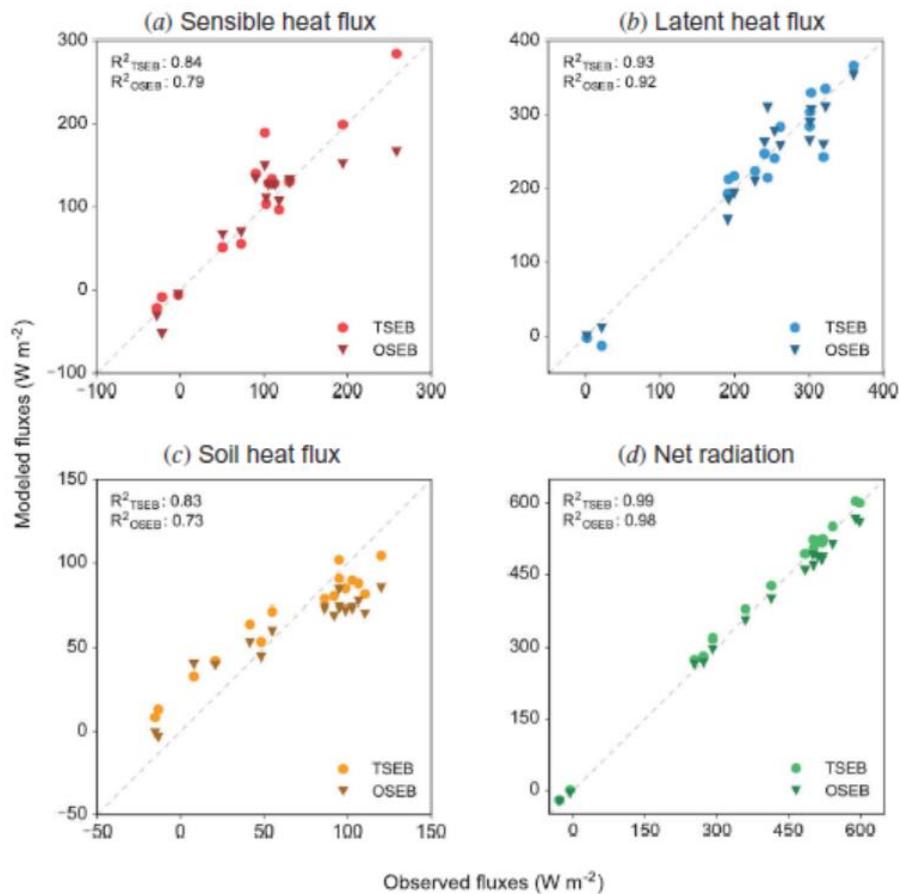


Figure 46: Comparison of modelled (based on UAV) and measured energy balance components (latent and sensible heat, ground heat flux, and net radiation; with EC). Fluxes modelled by the OSEB model are marked with circles, while the TSEB output is marked with triangles. OSEB/TSEB are one/two source energy balance models used to derive energy fluxes from the UAV LST -measurements as explained in Brenner et al. (2017).

Additionally, we applied the widely used Food and Agriculture Organization of the United Nations (FAO) Penman-Monteith (PM) formulation (Allen et al., 1998) and a modified PM approach called Surface Temperature Initiated Closure (STIC (Malick et al., 2016)). The fact that turbulent fluxes vary predominantly with solar radiation allowed us to look at how different variables at the surface and the lower atmosphere respond and whether they deviated from a linear response with variations of solar radiation. The analysis showed that the sensible and latent heat flux predominantly also vary linearly with solar radiation. The magnitude of the phase lags of both quantities were relatively small (several minutes) and changed with the prevailing moisture conditions. Surface energy balance residual approaches, which simulate the latent heat flux using the surface-to-air temperature gradient as input, were able to reproduce the reduction of the phase lag from wet to water limited conditions. Those approaches, which use the vapour pressure deficit as additional driver as the Penman-Monteith equation, showed significant deviations from the observed phase lags (Fig. 48 and 49), while the temperature gradient shows phase differences in agreement with the sensible heat flux and represented the wet-dry difference rather well. Our results suggest that temperature-gradient approaches may be

more suitable to estimate evapotranspiration from observational data in climates without substantial water stress. The reason is that the vertical temperature gradient contains relevant signals of soil moisture limitation as opposed to the vapour pressure deficit of the air. Schemes, which use vapour pressure deficit as additional input, require a dynamic, i.e., time-dependent, characterization of surface conductance to account for the strong phase lag in vapour pressure deficit (Renner et al., 2019).

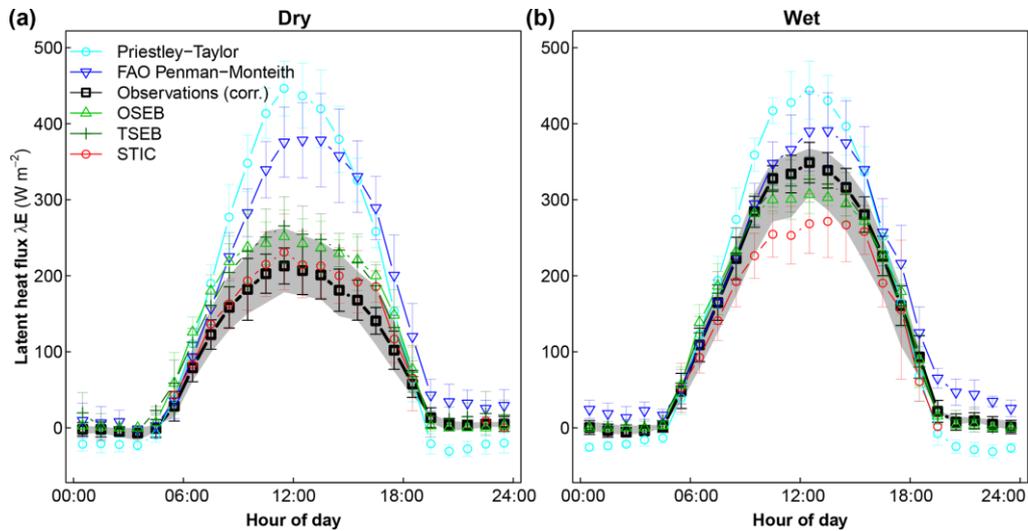


Figure 47: Average diurnal cycle of  $\lambda E$  estimates for (a) dry and (b) wet days. Error bars denote the standard deviation obtained for each hour. The bold black line with squares shows the observed latent heat flux corrected for the surface energy balance closure ( $\Delta EBRC$ ). The grey-shaded area depicts the range induced by the energy balance closure gap. Source: Renner et al. (2019).

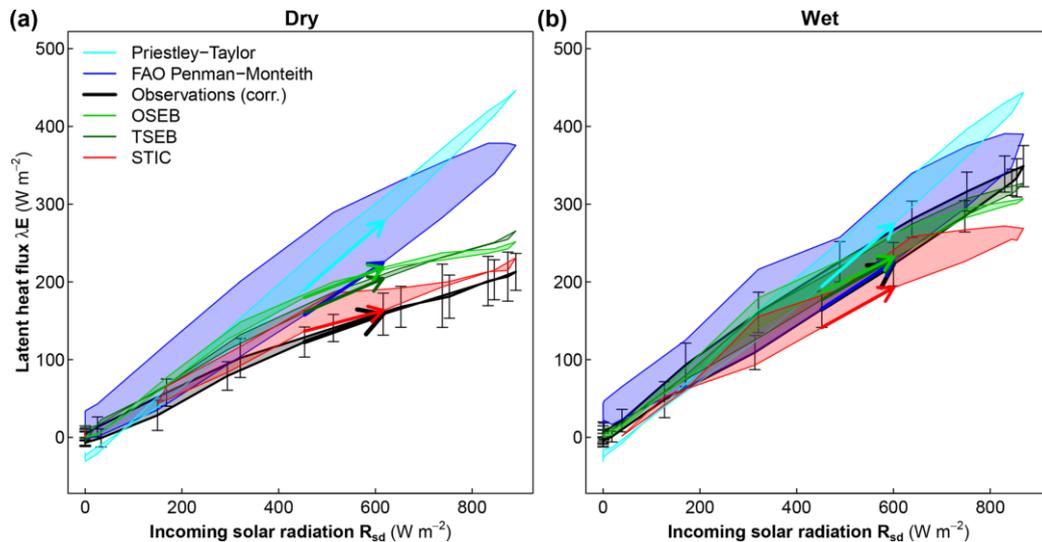


Figure 48: Diurnal hysteresis of  $\lambda E$  to  $R_{sd}$  for (a) dry and (b) wet conditions of observations and different models. Bold arrows indicate the rising limb in the morning hours (07:00 to 08:00 LT) showing a counterclockwise hysteresis of  $\lambda E$  under wet conditions. Vertical arrows depict the standard deviation of  $\Delta EBRC$  for each hour. Source Renner et al. (2019).

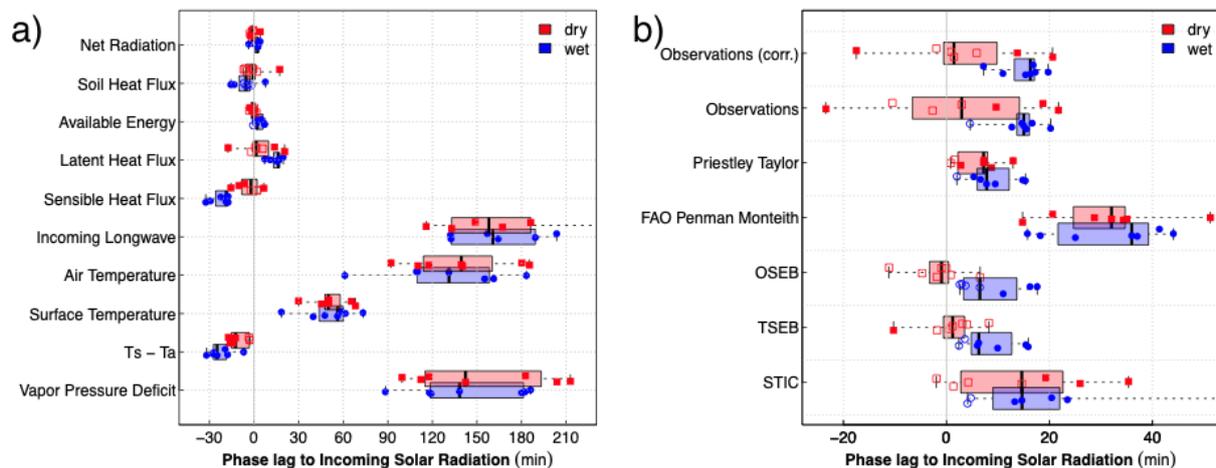


Figure 49: Phase lag to solar radiation of (left) surface energy fluxes and surface state variables used as input for the (right) evapotranspiration models for dry (red) and wet (blue) days. Boxplot and daily estimates with filled symbols showing significant phase lag estimates. From: Renner et al. (2019).

Phase I project D started to develop a completely new thermodynamic framework to understand and model land surface functioning, particularly regarding estimates of evapotranspiration and energy balance components. This motivated by the fact that the turbulent exchange of heat, water vapour and carbon dioxide between the land surface and the atmosphere requires physical power to accelerate air in the vertical direction. Thermodynamics adds an important constraint to the generation of power, which provides in concert with thermodynamic optimality principles a powerful means to explain and predict the radiation balance as well as portioning of net radiation into sensible and latent heat, as further explained in section 2.2.2.5.

#### 2.2.1.6 Synthesizing minimum data needs for the identification of functional units

A key question driving experimental research was about the minimum necessary sampling for to discriminate functional differences. Here we focus on the two most important descriptors of functional units namely geology and land use. To investigate whether we can reliably identify differences between functional units and how much data is necessary to achieve this we studied exemplarily a) event responses in the unsaturated zone and b) event responses in stream water quality. We specifically wanted to know at how many locations we need to measure to identify differences and if long-term measurements are necessary. In case a) we determined the fraction of events with a non-sequential moisture response (NSR) in the soil profile, as indicator for events where preferential flow occurred (see also chapter 2.2.1.3). We randomly sampled our pool of 135 soil moisture profiles, starting with a sample size of 1 and increasing it stepwise up to the full 135. Sampling was repeated 40 times for each sample size and the 5<sup>th</sup> and 95<sup>th</sup> percentile of the resulting distribution was plotted, showing how the distribution narrows with increasing number of N (Fig. 50, upper panel). Grouping our measurement locations reveals differences between both conditional distributions, and the minimum number of soil moisture profiles needed to discriminate these differences (Fig. 50, lower plot). While forests have a higher fraction of preferential flow responses than grassland, we need at least 18 profiles to reveal this difference.

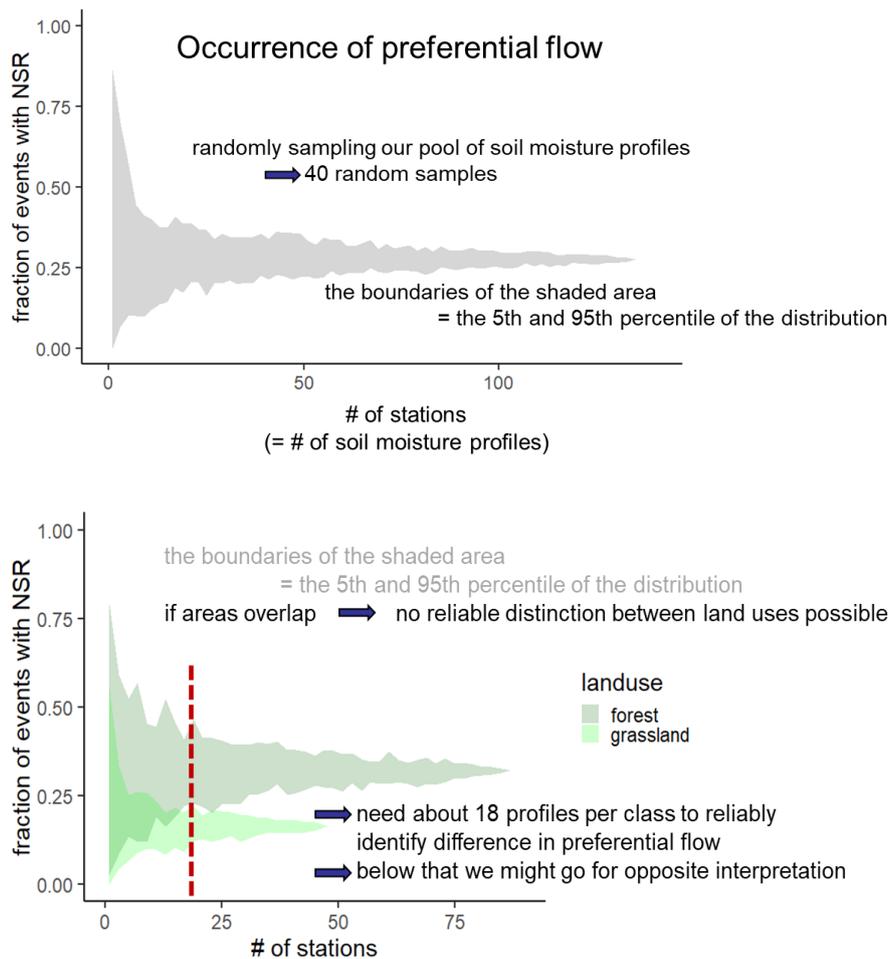


Figure 50: Distribution of the fraction of preferential flow events vs sample size (upper panel). Comparison of conditional distributions for forest and grassland sites and identification of minimum data sample size to discriminate having different fractions of preferential flow events with NSR. (Results to be published in Project G)

The next step was to look at conditional distributions of NSR events, using landuse and geology as confining variables. While this yielded similar results for Marls and Schists (Fig. 51, upper panel), in the case of the sandstone region we hardly have enough soil moisture profiles to make this distinction. In the case of the schist region, the difference between forest and grassland is very pronounced and we measured more soil moisture profiles than would have been necessary. We also identify seasonal differences in preferential flow response in all geologies, but again most pronounced in the schist region (Fig. 51, lower panel). Due to these pronounced differences only 10 soil moisture profiles are sufficient in schist, while in sandstone we need more than 20 to reliably identify that preferential flow occurs more often in summer than in winter.

Given the number of observed soil moisture profiles, no significant difference in preferential flow was found between the marl and the schist region, for both grassland and forest sites. A direct inter-comparison of preferential flow behavior the three geologies reveals that differences are more pronounced under forest than under grassland (Fig. 52, upper panels), where the number of observed soil moisture profiles were insufficient to discriminate differences among the geologies.

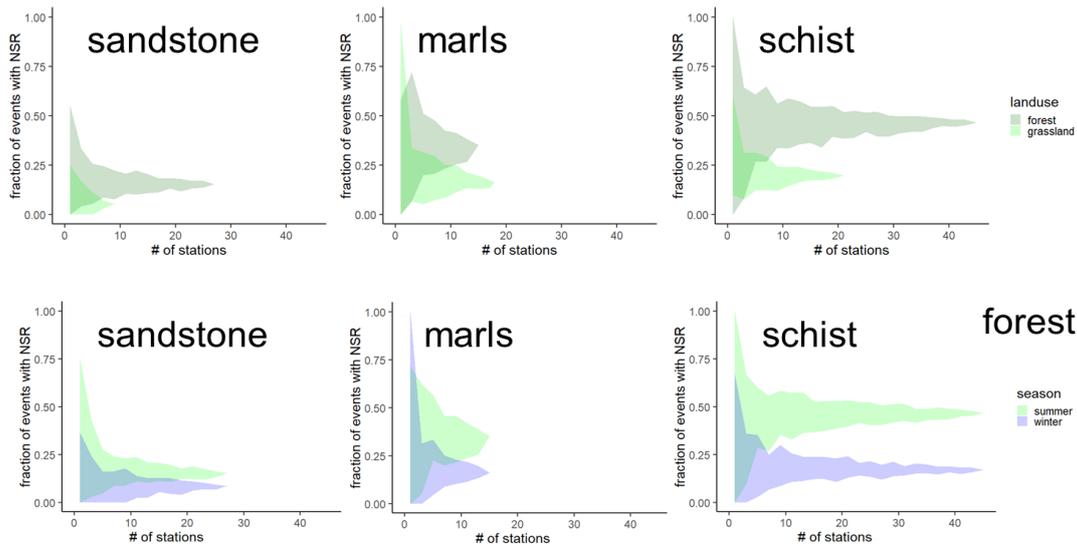


Figure 51: Comparison of preferential flow responses for forest and grassland across all geologies (upper panel). Seasonal differences in preferential flow responses under forest, grouped by geology (lower panel, results to be published in Project G).

To investigate the role of observation period on these differentiations, we reran the analysis for the forest sites using exclusively the data from 2016 (Figure 52, lower panel). With only one year of data it is no longer possible to differentiate between the preferential flow behavior of the marls, and the sandstone and the fraction of preferential flow under sandstone is overestimated. A shorter measurement period goes at the expense of a higher necessary amount of soil moisture profiles to reliably identify the differences between schist and sandstone (>10 for the full time series, 20 for the 1-year period).

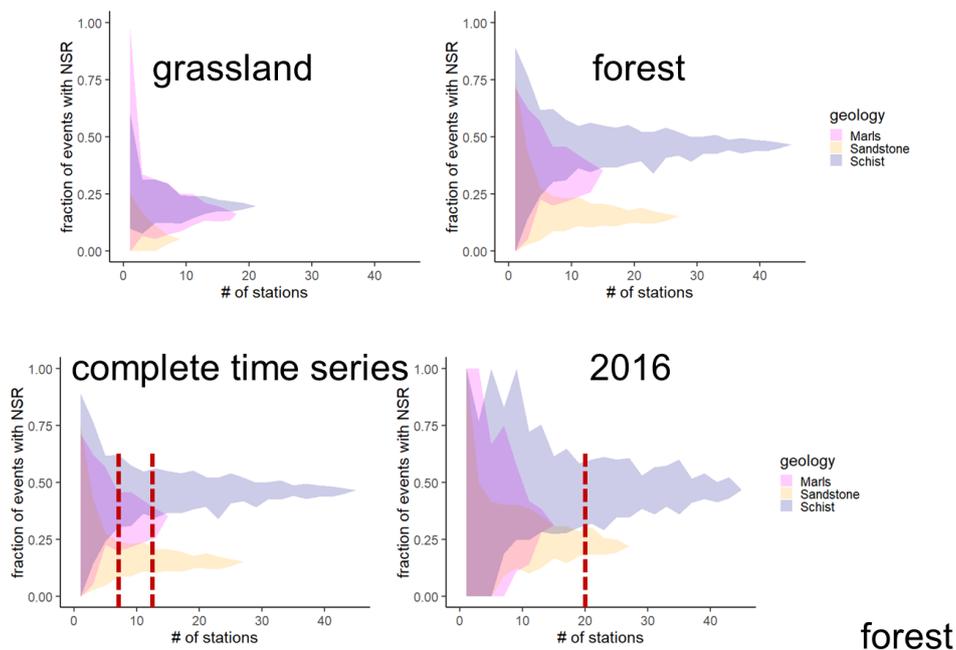


Figure 52: Differences in preferential flow response between geologies, grouped by land use (upper panel), lower panel: comparison of full data set vs. a single year.

This underpins the importance of long-term monitoring for reliable interpretations of the observed patterns, but also shows that a minimum of about 20 measurement locations (in this case soil moisture

profiles): per functional unit are required to identify differences and similarities between these functional units.

Using this same approach for the data set of stream water electric conductivity (see section 2.2.1.3.3) we investigated if there are differences in the fraction of rainfall events where the respective streamflow response produced a dilution effect with respect to stream solutes. Solute transport response to rainfall events can support indication of runoff generation processes but also contains information about mineralogy and anthropogenic effects. We found that the schist region behaves significantly different compared to the sandstone and marl regions (Fig. 53).

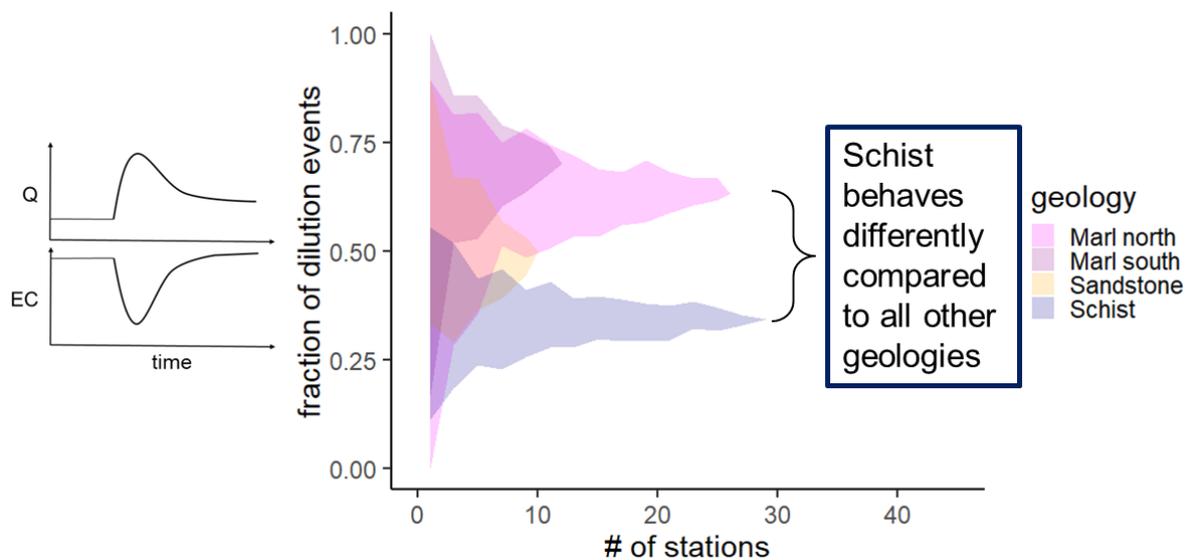


Figure 53: Comparing event responses based on stream solute dilution effect across the geologies (Results to be published in Project G)

A possible explanation for this is the rapid subsurface responses and subsurface connectivity in the schist (see also 2.2.1.3.3), compared to the more likely surface and tile drain runoff in the marls. However, electric conductivity in the schist streams is generally low, providing less of a difference to the EC of incoming rainfall. The fact that the sandstone behavior could not be differentiated from the marls might be due to the fact that runoff in the sandstone is actually generated on the impeding underlying marls layer and that some of the sandstone sites are only dominated by sandstone in the upper reaches, but by marls in the downstream and near-stream areas. While it would have been good to have more measurement locations in the sandstone, this was hampered by the very low stream density.

Dilution effects during streamflow response also differed strongly from summer to winter, across all geologies (Fig. 54). This is likely due to the fact that summer low flows have higher EC (especially in the marls) and thus a stronger difference to the incoming rainfall EC, but could also be linked to changes in runoff generation processes during the drier part of the year. Here stream temperature responses to rainfall events also hint at surface or near-surface runoff in the marls region (compare chapter 2.2.1.3.3).

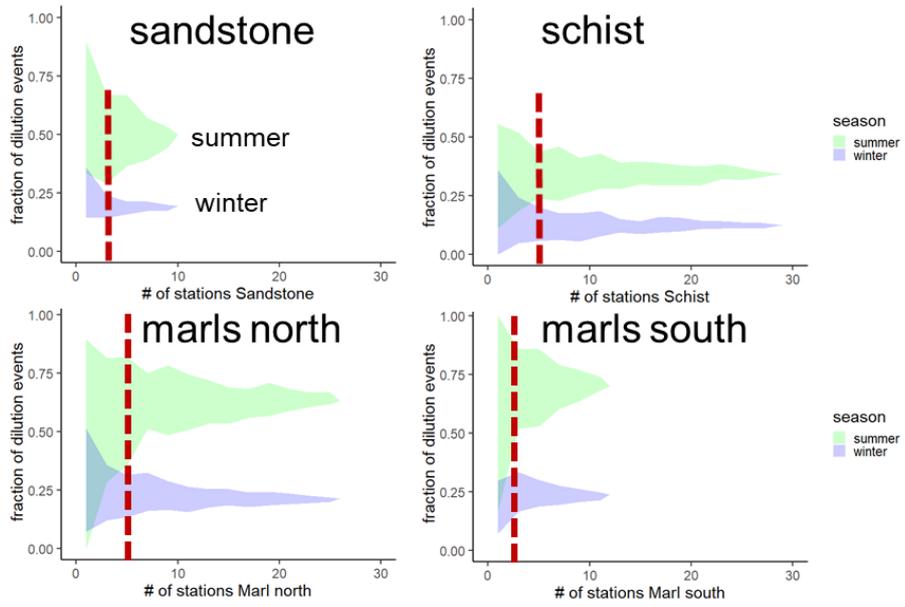


Figure 54: Comparing event responses based on stream solute dilution effect per season across the geologies (Results to be published in Project G)

## 2.2.2 Key model and theory developments

The previously presented experimental research went hand in hand with model development and theory building. Hydrological processes are on a very general level dissipative, as storage, mixing and release of water as runoff and latent heat are generally associated with conversions, dissipation and export of energy. A thermodynamic perspective appears thus as a natural choice to analyse the interplay of the mass and free energy balance of hydrological systems, explain its dependence on the prevailing forcing, system state and structure, and to explore related opportunities for constraining hydrological predictions. In phase II, we combined the thermodynamic approach with an information perspective. While the latter provides superior measures to quantify similarity, uncertainty and the degree of organisation of a system, the former is superior for constraining the joint mass and energy balance and for predictions. In the following, we shortly outline the major CAOS modelling exercises and relation to these theoretical concepts. These range from efforts to test:

- the central CAOS hypotheses H2 suggesting the existence of dynamic functional units and the feasibility of a dynamically changing, adaptive model complexity (H3),
- new model approaches for preferential flow and solute transport in soils,
- fully coupled simulations of land surface-atmosphere feedbacks,
- the value of a thermodynamic perspective for better explaining hydrological functioning and constraining related model predictions,
- the complementary benefit of an information perspective to diagnose similarity and organisation.

### 2.2.2.1 Functional units for runoff generation and dynamic hydrological similarity

#### 2.2.2.1.1 Modelling lower mesoscale catchments with representative hillslopes

Within two related model studies Loritz et al. (2017, 2018) corroborated that hydrological landscapes can be divided into hillslope scale functional units, which in turn can be used as building blocks for setting up simplified and yet physically sound hydrological models. The first study showed that the water balance of the forested Colpach in the Devonian schist/slate (19.4 km<sup>2</sup>) and the agricultural Wollefsbach in the Marls (4.5 km<sup>2</sup>), can successfully be simulated by a single 2d representative hillslope Fig. (55).

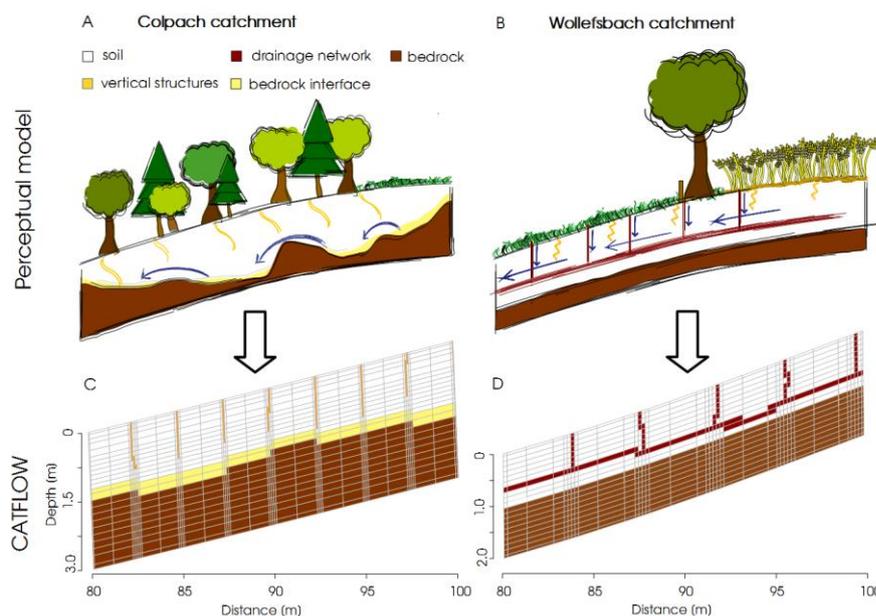


Figure 55: Perceptual models of the (a) Colpach and (b) Wollefsbach and their translation into a representative hillslope model in CATLFLOW. It is important to note that only small sections of the model hillslope are displayed (C Colpach; D Wollefsbach). Source Loritz et al. (2017).

These models were informed by the available comprehensive field data on soil hydraulic

properties (phase II project F), macropores (phase II project A) and geophysical data to constrain the bedrock topography (phase II project B group Tronicke). The corresponding simulations of stream flow (Colpach KGE 0.88 Fig. 56 a, Wollefsbach KGE 0.71), transpiration (Fig. 56 b) and distributed soil moisture dynamics were in both catchments in acceptable accordance with observations of the hydrological year 2013/14 (Loritz et al., 2017).

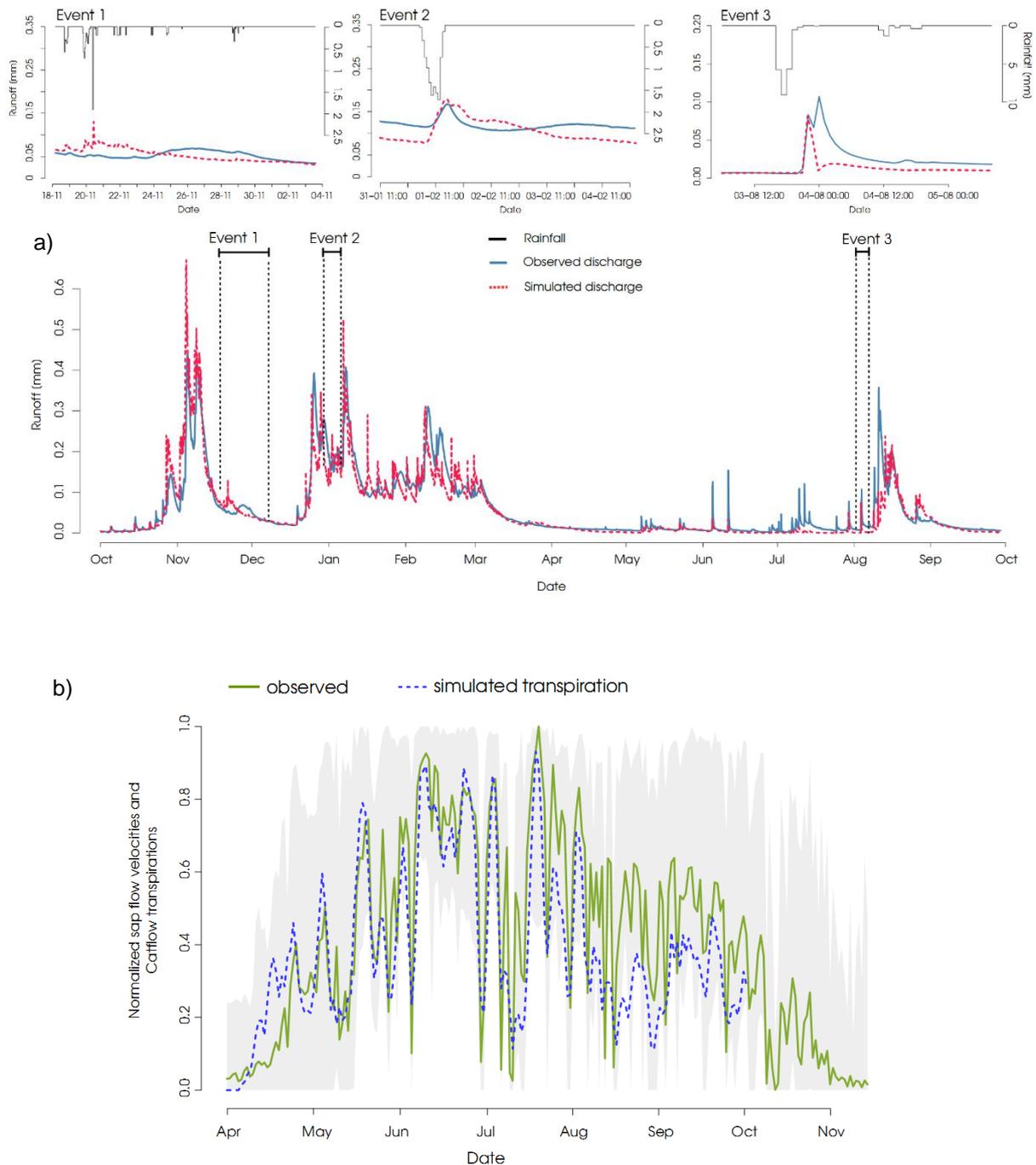


Figure 56: Observed and simulated runoff of the Colpach catchment in the hydrological year 2013/2014 and three selected events (a). Normalized observed average sap velocities of 28 trees in the Colpach catchment (green) and normalized simulated transpiration from the Colpach (dashed blue). Additionally, the ensemble of all 28 sap flow measurements is displayed in grey.

This underpins that the concept of representative hillslopes is a feasible approach to simplify distributed models without lumping, because they preserve the relevant information about the average driving gradients and resistance terms that control runoff generation and hydrological dynamics. The key to the derivation of these representative hillslopes was to respect energy conservation during the aggregation procedure. Specifically, we derived an effective topography such that it conserved the average distribution of potential energy along the averaged flow path length to the stream. Similarly, the macroscale effective soil water retention curve was constrained to preserve the relation between the average soil water content and matric potential energy using the set point scale retention experiments as explained in (Jackisch et al, 2017; Zehe et al., 2019, see also Fig. 4). Test simulations with randomly selected retention functions of individual experiments and based on the averages of the individual parameter sets performed clearly worse (Fig. 57).

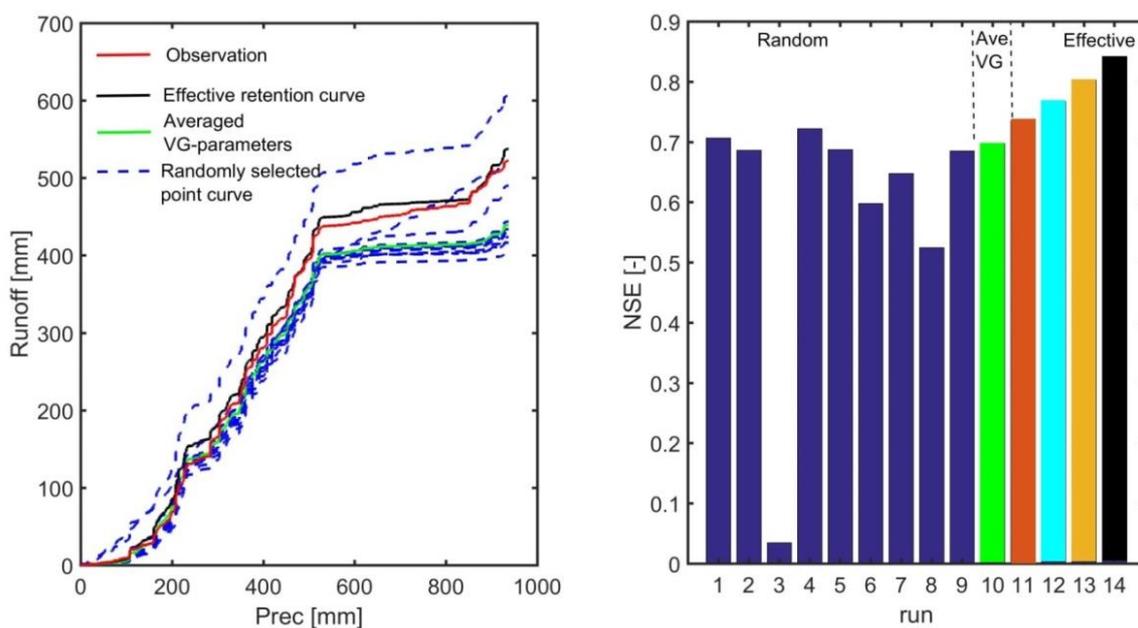


Figure 57: a) Sensitivity of simulated double mass curve to the choice of the retention function (accumulated specific runoff/streamflow plotted against accumulated precipitation in the hydrological year 13/14). B) Nash Sutcliffe Efficiency NSE for simulations based on randomly selected point curves (blue), the averaged van Genuchten parameters (green), the best model setup based on the effective retention curve (black), the same configuration without vertical or lateral macropores or any macropores (light brown, light blue, dark brown).

#### 2.2.2.1.2 The dynamic nature of functional similarity

To further illuminate the nature of hydrological similarity, Loritz et al. (2018) analyzed simulations of a fully distributed setup of the Colpach catchment consisting of 105 hillslopes with variable topography and identical subsurface structure using the macroscale effective soil water characteristics mentioned above (Fig. 58). By means of the Shannon entropy (Shannon, 1948) we quantified the diversity in simulated runoff and storage of the hillslope ensemble at each time step (Fig. 59 a and b). Note that an entropy maximum implies that each hillslope contributes in a unique fashion, while a value of zero implies that all hillslope yield a similar runoff response. This revealed that the entropy of the runoff ensemble was rather dynamic in time, while it never reached the maximum value, which implied that the hillslopes yielded strongly redundant runoff contributions to stream flow (Fig. 59 c).

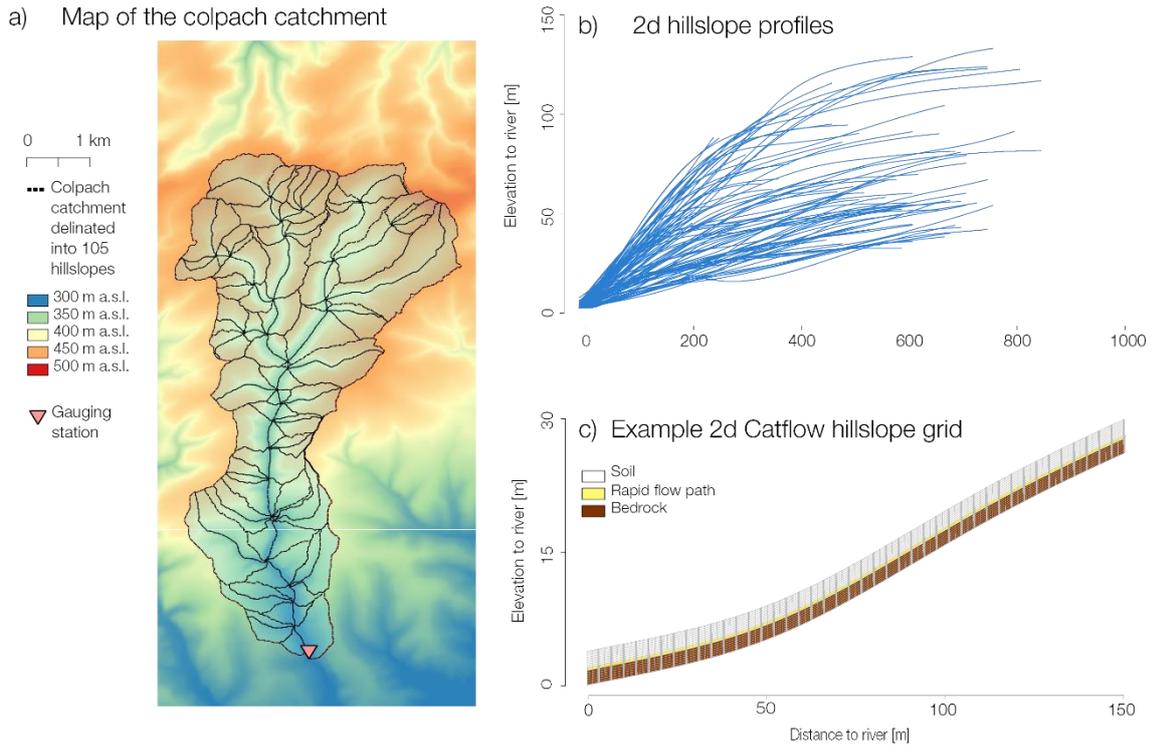


Figure 58: a) digital elevation model of the Colpach catchment and its delineation into 105 hillslopes b) all hillslope profiles extracted using the LUMP approach c) example of a CATFLOW hillslope grid. Source Loritz et al. (2018).

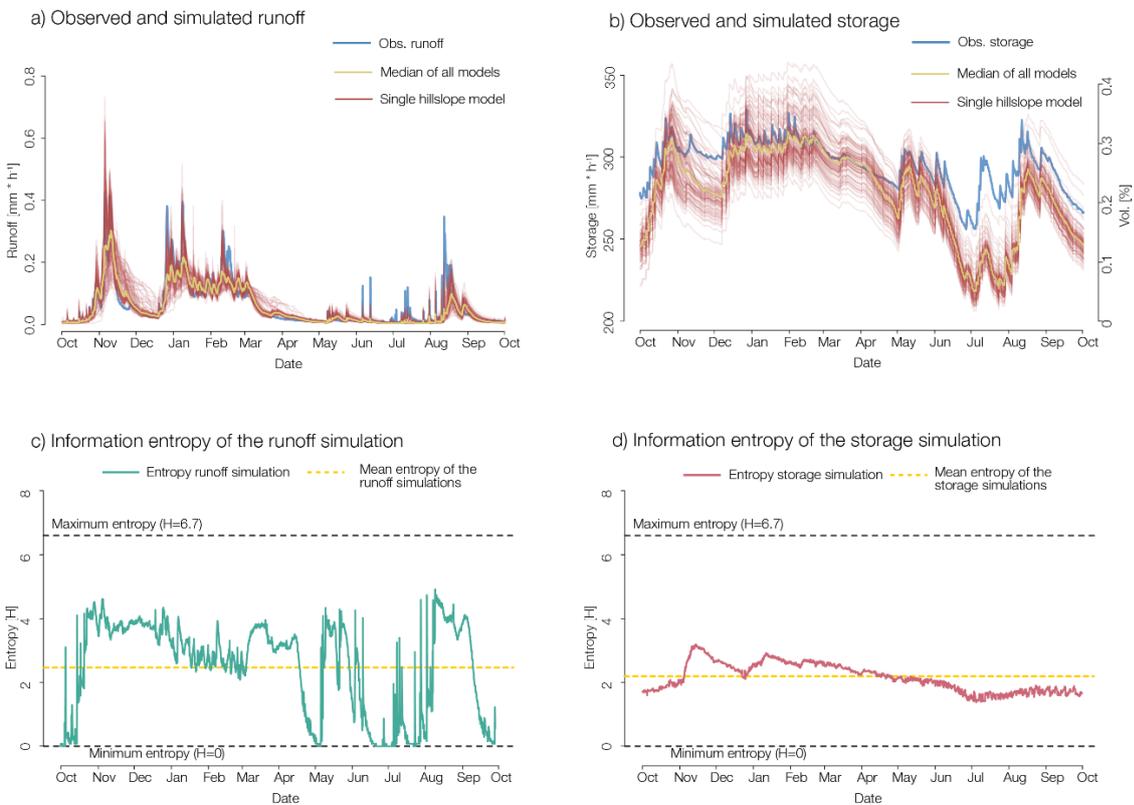


Figure 59: (a) Observed and simulated runoff of the Colpach catchment. (b) Simulated total area specific storage of each hillslope, the medians of all models and of observed storage. (c) Shannon entropy of simulated runoff and storage (d) and corresponding averaged entropy. Source: Loritz et al. (2018).

The entropy of total storage exhibited a smaller temporal variability and was smaller, suggesting a lower sensitivity of storage to variations in topography (Fig. 59 c). We further showed that the set of 105 hillslopes could be clustered into six functional units of similar runoff response based on their mutual information (Fig. 60). When using scaled runoff simulations from six arbitrarily picked representatives of these functional units, their sum performed on average as good to the full ensemble as further detailed in Loritz et al. (2018). We thus conclude that the fully distributed model can be strongly compressed without information and thus performance loss, thereby avoid redundant computations.

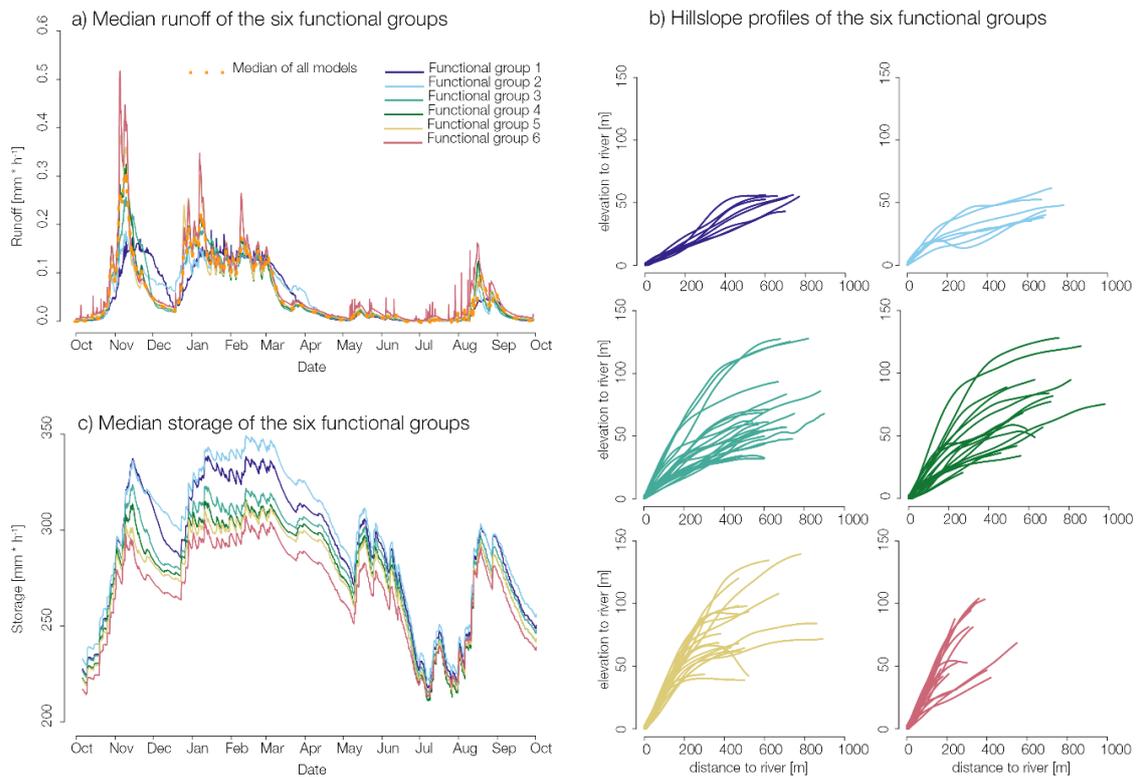


Figure 60: Median runoff time series of the six functional units; b1 –b6 (a), corresponding hillslope profiles with the elevation to river on the y axis and distance to river on the x axis for each functional group(b) as well as median storage time series of the six functional groups (c). Source: Loritz et al. (2018).

This finding corroborates the existence of functional units for runoff generation and it furthermore explains why conceptual models with 5-6 parameters simulate rainfall runoff behavior of catchment often very well, as most of the topographical heterogeneity in the catchment is irrelevant for runoff production. Recent studies yielded similar findings for a simple conceptual model (Ehret et al., 2020) and for CATFLOW when being used with gridded radar based rainfall estimates (Loritz et al., 2021), as further detailed in section 2.2.2.2.

### 2.2.2.2 Adaptive models for the water and energy cycle

#### 2.2.2.2.1 The CAOS model – an adaptive model for the water and energy balance of lower mesoscale catchments

A central objective of phase I project S and later phase II project E was to develop a process-based model for water-, energy-, and mass flows for lower mesoscale catchments using elementary functional units (EFUs) as building blocks. Key emphasis was on an explicit representation of networks controlling

preferential flow such as macropores, pipes, or rills. Scope of this development was to create a model framework, which allows testing of the central CAOS hypotheses, i.e. feasibility of an adaptive grouping of EFUs into cluster of similar functioning and to represent those by a single of few representatives thereby avoid redundant simulations. This was deemed to step beyond the static model compression presented in the previous section and to reduce computational loads such that the model becomes applicable in engineering practice without too much compromising the physics governing the partially zone (preferential) flow processes. In phase I a first version of the CAOS model was implemented in Matlab in a fully object oriented manner. Here we put major emphasis on the physical soundness of governing equations, the explicit treatment of vertical and lateral preferential flow in connected networks and a technically sound and efficient numerical solution of the model equations (Zehe et al., 2014). The model consisted of object hierarchy closely resembling real physical landscape elements (Fig. 62), with the catchment object on top, followed by hillslope and riparian zone objects which are further divided into Elementary Functional Units (EFUs), which are assumed to be laterally homogeneous (ca. 1.000 m<sup>2</sup>), but further subdivided in layers. Model execution followed the hierarchical model setup, with variable time stepping for within-objects; fluxes across object boundaries are updated at a predefined global time step.

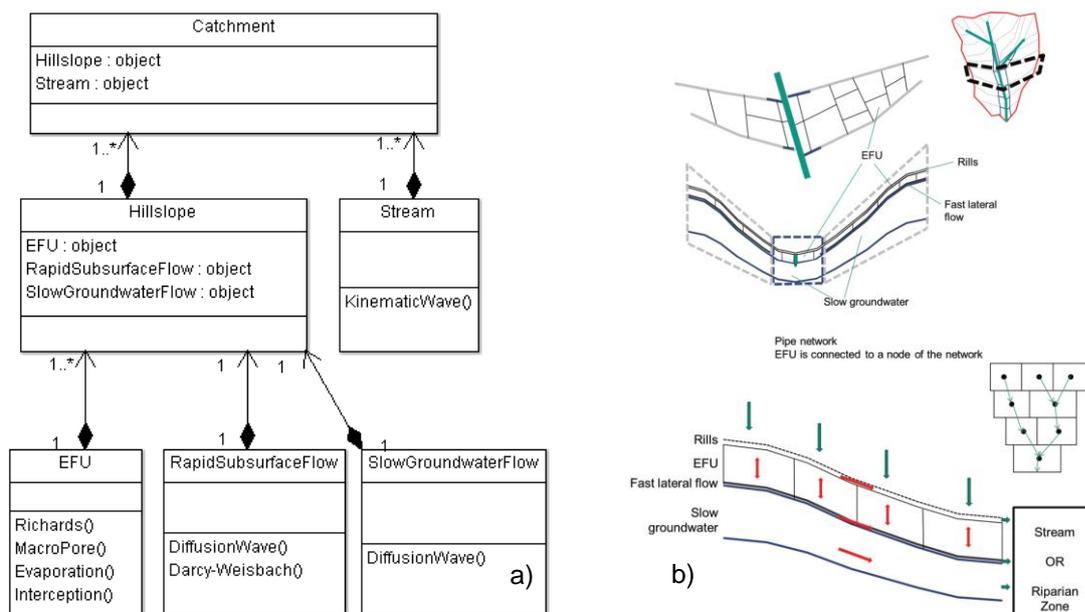


Figure 61: a) Simplified UML diagram of the model structure of the first CAOS model version. Each object either has child objects or solves 1-d flux –balance equations. B) Holzriver catchment and sketch of the model elements.

The EFU object resolves the flow in the unsaturated domain, including evaporation and transpiration. The core of this object is the soil matrix domain, which was solved with the Richards equations in the mass conservative ‘ $\theta$ -based’ version and Picard iteration. Water beyond saturation is directed either to the macropore domain (infiltration excess) or to the Rapid Subsurface Flow object. The lower boundary condition is free drainage, which connects to the Slow Groundwater Flow object. Soil evaporation is subtracted from the top grid cell, while transpiration is subtracted from all grid cells within the rooting depth. Transpiration is determined with the Penman approach, where resistance terms depend on water availability and wind speed, as implemented in the CATFLOW model. The macropore domain is modeled as a non-linear reservoir with macropore densities obtained from small scale sprinkling

experiments and ecological exploration (compare section 2.2.1.2). The rapid subsurface flow domain receives its water from the all EFUs within the hillslope and is represented as convergent pipe flow network. Flow within the network is, depending on the flow situation, modeled with either the diffusion wave approach or the Darcy-Weisbach equation. The slow groundwater domain represents hillslope baseflow and is modeled with the diffusion wave equation. It receives its water from the lower boundary of the matrix domain and the rapid subsurface flow object. Groundwater flow on the hillslope is assumed to be laterally homogeneous. The stream domain is represented as open channel network, receiving its water from the Rapid Subsurface Flow and Slow Groundwater Flow objects of all connected hillslopes. Flow is described with the kinematic wave equation and solved with an adaptive time stepping Crank-Nicolson-scheme. Each of CAOS model objects may account for reactive transport of solutes or heat based on the advection-dispersion equation, which is also solved with an adaptive time stepping and the Crank-Nicolson scheme. With respect to model complexity, this CAOS model version steps as originally intended clearly beyond the REW concept (Reggiani et al. 2005, Lee et al. 2007) as it avoids averaging across landscape components of different function and hence allows closure of the mass, momentum and energy balance in a spatially resolved manner. On the other hand, the model is clearly simpler as fully distributed, physically based models as for instance HydroGeoSphere (Brunner and Simmons 2012) or CATFLOW.

Despite of the soundness and the computational efficiency of this CAOS model version, the object structure did not allow for a quantification of energy and information flows associated the simulated hydrological dynamics on runtime. As this is indispensable for the implementation of adaptive model structure and the related quantification of dynamical grouping of functionally similar model objects (see next section), phase II project E started a complete new development of a second CAOS model version. This effort put much more emphasis on a sufficiently flexible object structure and the possibility to track mass, information and energy fluxes in the model. This went on the expense of a simple numerical treatment of the model equation using an explicit Euler forward scheme. The second CAOS model version is fully object-oriented to represent water-related processes in the atmospheric boundary layer, the soil surface, unsaturated and saturated zone and vegetation. The model is coded in Matlab with full versioning control in GitHub. In this framework, so far all but the boundary layer and vegetation processes were implemented; these are still represented by observations. The model represents 3d processes by multi-1d representations to reduce computational efforts and explicitly represents all stocks and fluxes of mass (surface water flow, matrix and preferential vertical infiltration, subsurface lateral flow in preferential flow structures, saturated flow in alluvial floodplain, river flow) and energy (thermal energy, geopotential energy, capillary potential energy, kinetic energy). All structural and state variables in the model are physical quantities, i.e. they can be directly estimated from observables. The main model structure and elements are shown in Figure 62. A model-setup exists for the entire Attert basin, which was used to generate high-resolution (5 min, few m<sup>2</sup>) model runs for several sub basins. We also developed an advanced GIS- and Matlab-based preprocessing tool for model setup in new catchments. Despite of these efforts the model is far from being complete as it misses process representations related to the boundary layer and soil-atmosphere exchange and evapotranspiration. The reason for these misses are the challenges of formulating all processes in the model with a full representation of all related stocks and fluxes of mass and energy, which were underestimated, as well

as considerable difficulties due to high computation times and numerical errors, due the very simple numerical scheme.

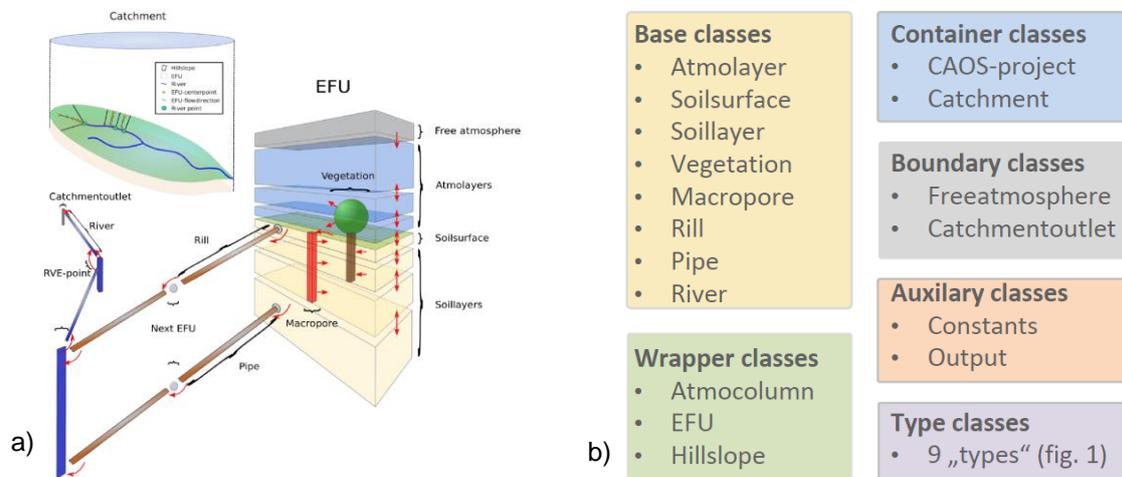


Figure 62: CAOS model structure (a) and model classes (b).

We must thus admit that we so far clearly didn't live up to our goal to deliver a model of intermediate and adaptive complexity that allows feasible simulations of the water and energy balance of lower mesoscale catchments. Despite to this drawback the efforts were valuable as they stimulated a fruitful exchange with other teams of the CAOS research unit. The feasibility of adaptive modelling was, nevertheless, tested using two essentially different model structures (Ehret et al., 2020; Lortz et al., 2021) as detailed in the next section.

#### 2.2.2.2.2 Adaptive model clustering: improving modelling efficiency by using dynamic functional similarity

Project E designed and coded a new conceptual model called SHM (Simple Hydrological Model) as a testbed for dynamical clustering and applied it to the Attert basin. The main goal of adaptive clustering is to reduce computational efforts of distributed and high-resolution modelling to facilitate application at larger scales or for longer periods. The main idea is to avoid redundant computations by clustering similar model elements, and then to infer the dynamics of all elements in a cluster from just a few representatives. In contrary to the above presented static compression of the 105 CATLFLOW hillslopes into fixed number of six functional groups, adaptive clustering relies on dynamic number of functional units. For adaptive clustering to make sense, three preconditions must be fulfilled: Existence of i) many model elements of ii) the 'same kind', with potentially similar behavior but iii) only weak interaction. These preconditions are largely fulfilled for sub catchments or hillslopes in distributed hydrological models. They occur in large numbers, there is only little or no interaction among them as they act in parallel, connecting to rivers, and as the critical zone in the landscape is composed of relatively few, typical, recurring combinations of its constituents there is potential for similarity among model elements, which we can exploit. It is important to note that even if two model elements were identical with respect to all of their time invariant (structural) properties, they could still behave differently when exposed to different forcing. Along these lines Ehret et al. (2020) developed ways to analyse dynamical similarity of model elements and re-cluster them during runtime using concepts from information theory, as briefly explained in the following (Fig 63).

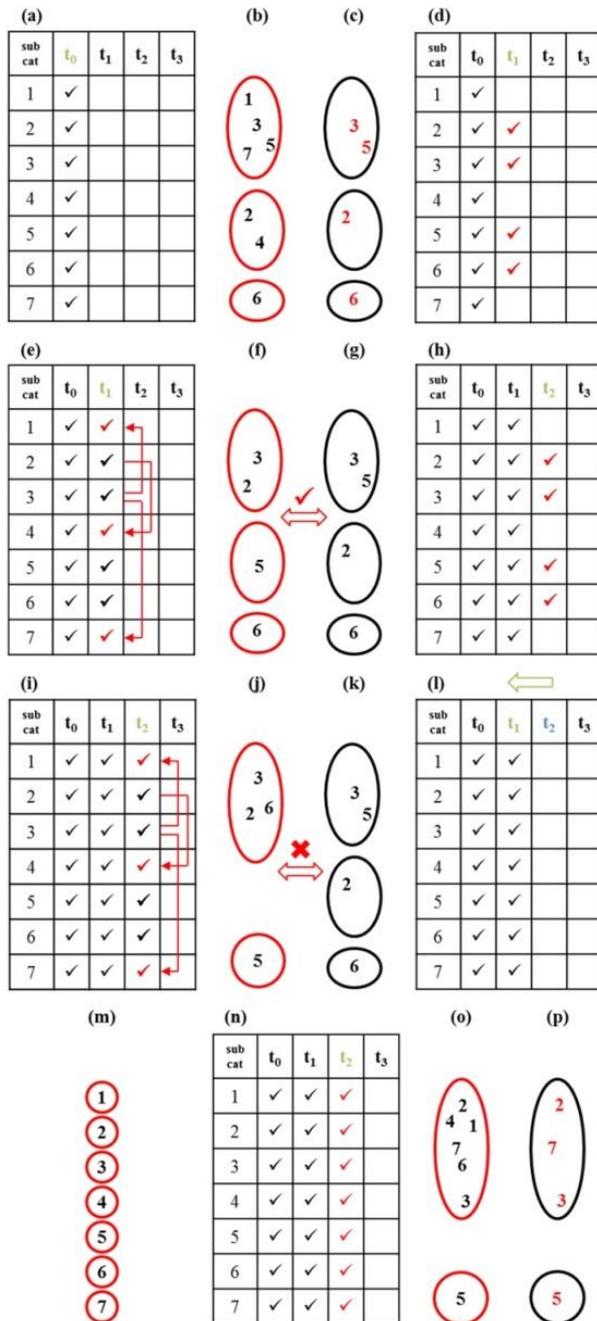


Fig 63. Main steps of adaptive clustering. A detailed explanation is given in the following text.

In the matrix shown in Fig. 63 a rows represent model elements (sub catchment), and each columns a time step. Starting from a fully distributed initial state at  $t_0$ , each sub catchment is simulated separately to determine its own states and fluxes. Based on this, objects are clustered in dynamically similar sub catchments in groups (Fig. 63 b) using i) a suitable clustering algorithm and ii) choice of one or several sub catchment properties acting as control variable for the clustering. Next, a subset of representative elements is selected from each cluster (Fig. 63 c). Their number and their choice strongly influences the performance of the adaptive model. A large number favors a high simulation quality on the expense of a small computational gain and vice versa. Generally, the choice of the control parameters aims to minimize computational expense at a minimum deterioration of simulation quality compared to the fully distributed model. Once the representatives are picked, the next time step  $t_1$  is simulated and done for those (Fig. 63 d); the corresponding states and fluxes are assigned to the remaining member of the sub catchments clusters without compromising mass conservation (Fig. 63 e). The larger difference between the computational costs of

this mapping and execution of fully resolve model, the larger is the computational benefit of the clustering procedure. As there is no guarantee that the initial clustering, based on control variables at  $t_0$ , will be still be valid at the new time step  $t_1$ ; a new clustering is done based on the values of the clustering control variable at time step  $t_1$ , but only for the chosen representatives (Fig. 63 f). Based on the similarity of this new clustering to the initial one (Fig. 63 f-g), the initial clustering either still holds and the previous steps are repeated (Fig. 63 d-e) for the next time step  $t_2$  (Fig. 63 h-i) or a new clustering is done, jumping back in time (Fig. 63 l-p). Expressing similarity of model elements and input data by their information entropy revealed striking patterns of spatial redundancy (Fig. 64 a and b) and seasonal and event-specific variations (Fig. 64 c). One of the key findings in this context was that prior knowledge of these patterns can greatly reduce computation times of spatially distributed models while the related decrease of simulation quality is in most cases acceptable (Fig. 65).

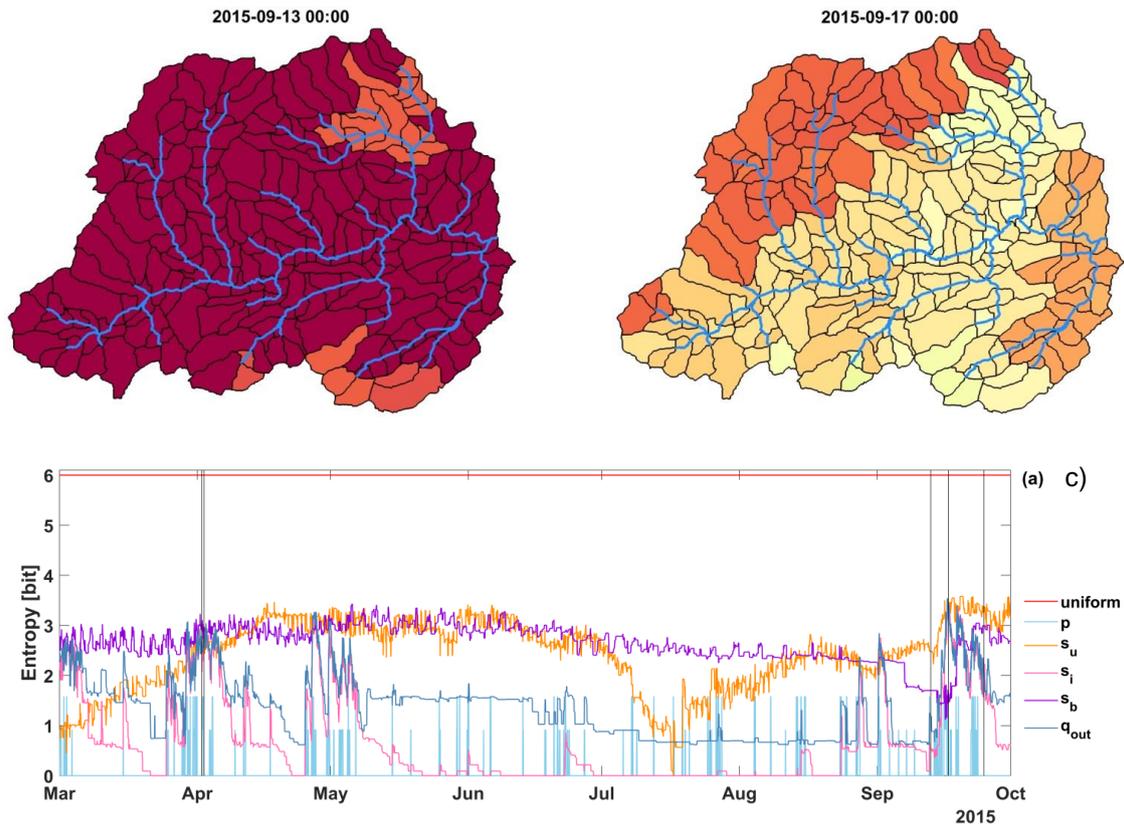


Figure 64: Spatial pattern of SHM Attert sub catchment outflow for two times. After a long dry spell, almost all sub catchments show the same low outflow (ab). After spatially distributed rainfall, a mixed pattern controlled by geology and spatial rainfall distribution appears (b, see also Fig. 2). Time series of Entropy of forcing ( $p$  = precipitation), states ( $s_u$  = unsaturated zone storage,  $s_i$  = interflow reservoir storage,  $s_b$  = base flow reservoir storage) and fluxes ( $q_{out}$  = sub catchment outflow) of all 173 sub catchments in the SHM Attert model (c). Source: Ehret et al. (2020).

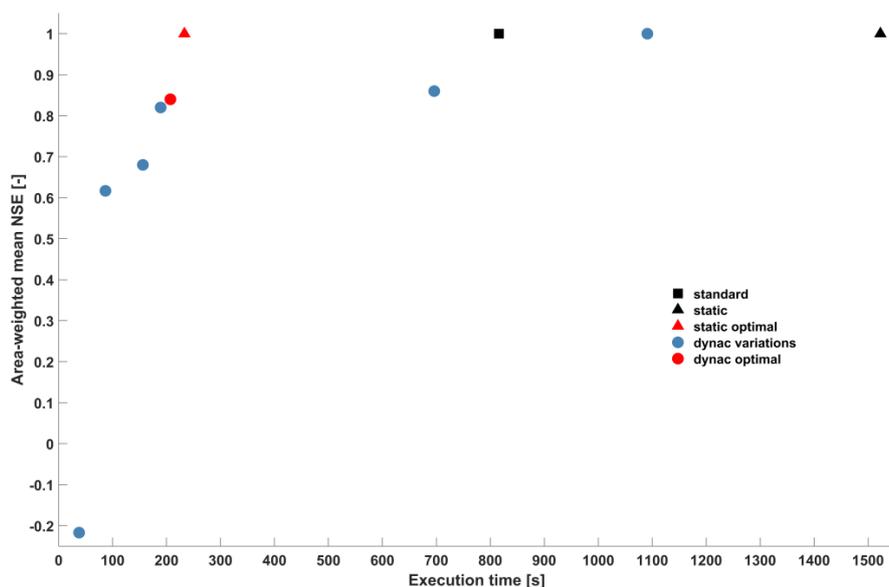


Figure 65: Variations of the SHM Attert model in terms of efficiency expressed by execution time, and effectiveness expressed by Nash-Sutcliffe efficiency. Compared to a standard full resolution run ('standard'), static optimal clustering based on prior analysis ('static optimal'), and dynamical clustering done during model execution ('dynac variations', 'dynac optimal') greatly reduces execution time while NSE deteriorates only to a lesser degree. Source: Ehret et al. (2020).

Even for this 'on-the-fly' clustering, execution acceleration by a factor of four at acceptable model simulation qualities could be achieved. For further technical details, please refer to Ehret et al. (2020).

Along similar lines but with emphasis on the role of spatially gridded precipitation for runoff simulations, Loritz et al. (2021) explored the feasibility of adaptive modelling using a set of 42 hillslopes in CATFLOW. These were structurally identical to the aforementioned representative hillslope model of the Colpach and driven by the gridded, radar based precipitation field shown in Fig. 38. This study revealed that distributed precipitation did generally clearly improve stream flow simulations compared to simulations with spatially uniform precipitation, particularly during the summer season, where frequently localised convective events occur. For two selected rainfall runoff events Loritz et al. (2021) showed that adaptive modelling is also feasible when using so called physically based models (Fig. 66). This may be explained by the dissipative nature of runoff generation, which implies that the memory of the system on distributed forcing and systems states is a rather short one (Fig. 67). This makes hydrological systems distinctly different from the atmosphere, which is susceptible to deterministic chaos.

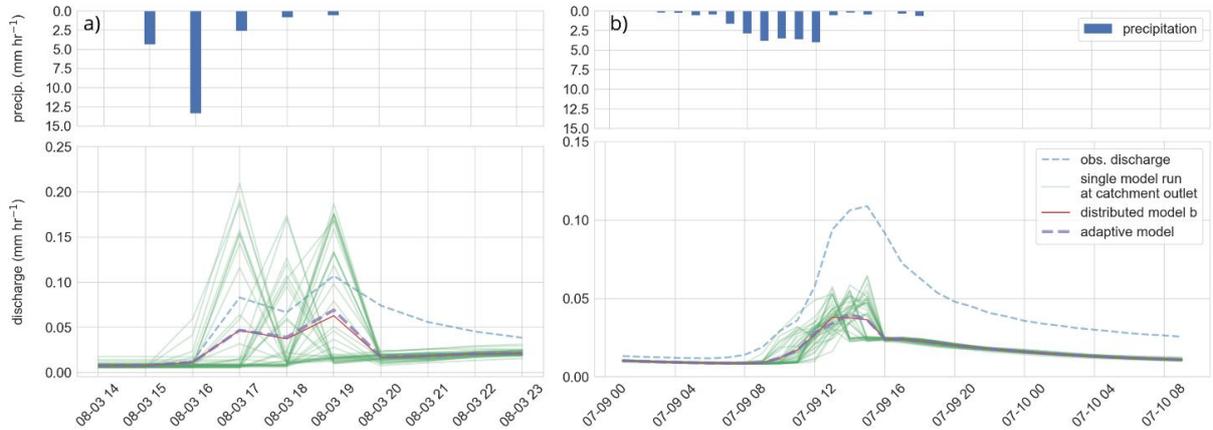


Figure 66: Rainfall-runoff event I (a) and rainfall-runoff event II (b). Blue bars in the upper panel show the average precipitation of the precipitation field for each time step (mm hr<sup>-1</sup>). The green curves in the lower panel represent a single gridded model of the distributed model b; red line the area-weighted mean of the distributed model, purple dashed line the area-weighted mean of the adaptive model and dashed blue line the observed specific discharge of the Colpach. Source: Loritz et al. (2021).

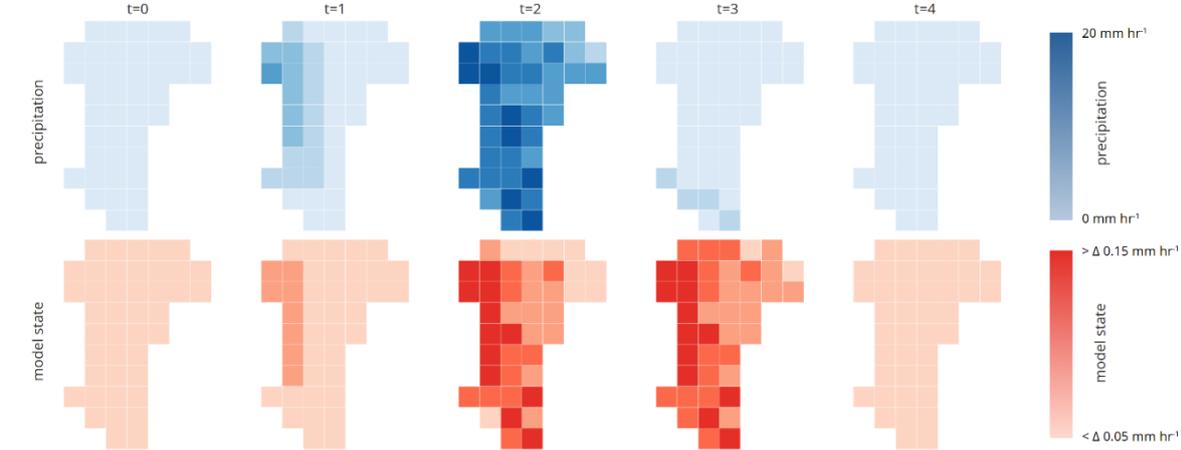


Figure 67: Spatial and temporal distribution of the precipitation field (upper panels) and the corresponding states of the corresponding model hillslopes as function of time. Source: Loritz et al. (2021).

### 2.2.2.3 New model approaches for preferential flow

#### 2.2.2.3.1 A Lagrangian framework for soil water dynamics and preferential solute transport

The inability of the Richards equation to account for imperfect mixing and preferential flow in structured soils has been known since a long time (Beven and Germann, 1982). Although a range of approaches has been proposed to address this problem (Šimůnek et al., 2003; Beven and Germann, 2013), none of those is commonly accepted as superior. Phase I project I started to develop a Lagrangian framework for simulating soil water dynamics and later on, in phase II project F, solute transport in structured heterogeneous soils. The Lagrangian perspective has been up to now used to simulate advective-dispersive transport of solutes (Berkowitz et al., 2006). Lagrangian descriptions of the fluid dynamics itself are only realised in a few models namely SAMP (Ewen, 1996a, b), MIPs (Davies et al., 2013).

Zehe and Jackisch (2016) conceptualized the first version of a Lagrangian model describing soil water flow using a non-linear space domain random walk, motivated by the equivalence of the Fokker Planck equation and the diffusive form of the Richards equation. In line with Ewen (1996), their model estimates the diffusivity and the gravity-driven drift term of the random walk based on the soil water retention curve and the soil hydraulic conductivity curve. A naive random walk, which assumes all water particles to move at the same drift velocity and diffusivity, overestimated depletion of soil moisture gradients compared to a 1d Richards solver (Fig. 68). The solution to this was to account for variable velocities, as characterized by the shape of the soil hydraulic conductivity curve (Fig. 69).

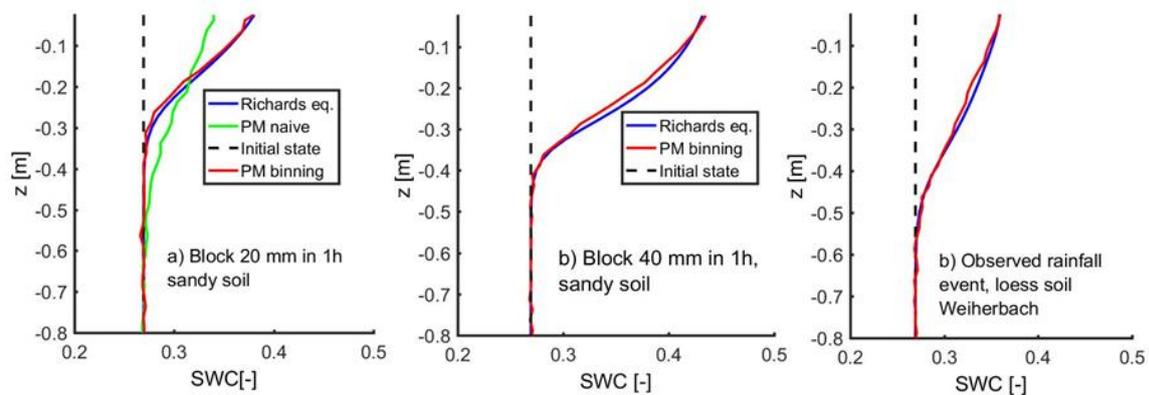


Figure 68: Soil moisture profiles simulated for a) a sandy soil and a block rain of 20 mm with a naive random walk (PM naive) and the Lagrangian particle model with binning (PM) according to Eq. 5 compared to a simulation with a Richards model, b) a sandy soil and a block rain of 40 mm, and c) an observed convective rainfall in the Weiherbach catchment. Source: Zehe and Jackisch (2016).

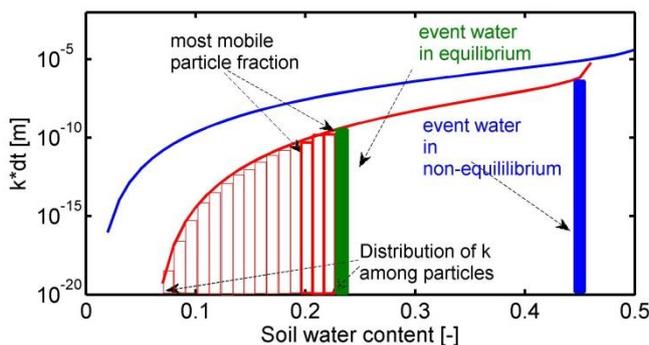


Figure 69: Advective displacement of a particle within a time step related to soil water content to illustrate the binning concept (Zehe and Jackisch, 2016) and the possibility to distinguish event and pre-event water.

Depending on the actual water content, the travel velocities of water particles are distributed accordingly using a suitable

binning. Because of this, the Lagrangian model accounts per default for the distribution of flow velocity within different pore size of soil matrix. It allows furthermore for a separated treatment of pre-event and event-water particles and non-equilibrium infiltration as detailed in Zehe and Jackisch (2016).

This one dimensional approach was expanded by Jackisch and Zehe (2018) into a two dimensions using a structured domain containing macropores as one vertical dimensional explicit structures. Within those, the velocity of each particle is described by interactions of driving and hindering forces using a generalized Bernoulli equation (Fig. 70) following thermodynamic reasoning.

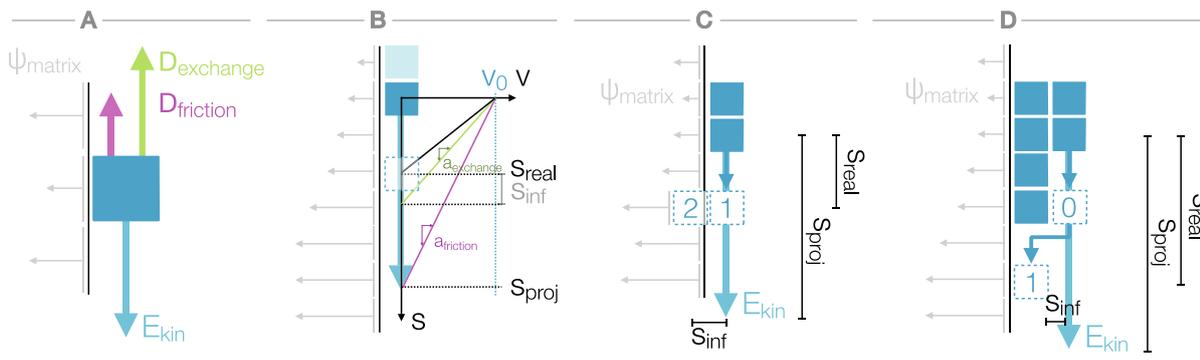


Figure 70: Macropore flow concept. (A)  $E_{kin}$  is dissipated by frictional and capillary forces at the macropore wall, (B) Projected advection with  $v_0$  is decelerated, (C) possible infiltration if contact time exceeds infiltration time, (D) Fast advection of a particle as film flow until the end of the film.

The driver is the geopotential energy of a particle, while energy dissipation occurs due to frictional and capillary forces at the macropore walls. The assets of this new echoRD model are a (i) self-limiting film flow in macropores, (ii), the ability to simulate 2-D infiltration patterns based on (iii) observable parameters. The model has been proven capable to simulate non-uniform infiltration patterns in accordance with observed patterns and to explore related structural controls on travel distance distributions (Jackisch and Zehe, 2018). The model was tested within different dynamic macropore settings in close cooperation with phase II project A (Reck et al., 2018, see 2.2.1.2.1). The results confirm the feasibility of echoRD to reproduce observed infiltration patterns (Figure 71).

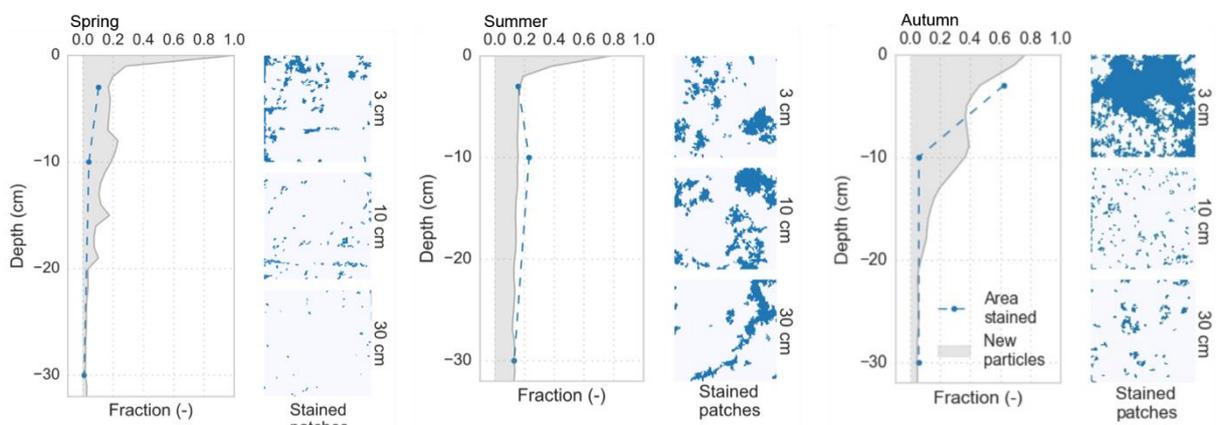


Figure 71: Observed staining in infiltration experiments at different seasons (blue) and modelled infiltration without calibration (grey). Source: Reck et al. (2018).

Sternagel et al. (2019) generalized the Lagrangian model to allow for simulations of water and solute transport in the matrix and a simplified separated macropore domain. Water particles are characterized

by their location, mass and solute concentration. A size and depth distribution and hydraulic radius characterize the macropore domain. The latter is helpful to determine the fraction of particles having contact to the macropore-matrix interface. Flow in the macropore domain is purely advective and the related velocities can be inferred either from tracer data or from direct observations reported in the literature (Zehe and Blöschl, 2004; Weiler, 2001; Shiptialo and Butt, 1999). The exchange water flux of pre-event water particles in the macropore domain with the surrounding matrix is calculated according to Darcy's law. Infiltration into the matrix and the macropores depends on their moisture state. A comparison of Hydrus 1d and the LAST-Model based on plot scale tracer experiments in the Weiherbach catchment revealed that both models show a similar good performance in case of matrix-flow-dominated tracer transport (Fig. 72, Hydrus 1d simulations are not shown). In case of preferential transport LAST clearly outperformed Hydrus 1d (Sternagel et al., 2019). The macropore domain for this site was parameterized based on a local survey of the worm burrow system similar to those carried out in phase II project A. The corresponding flow velocities were calculated based on their cross sections, using a regression derived from flow experiments with undisturbed soil samples containing macropores of variable diameters (Zehe and Blöschl, 2004).

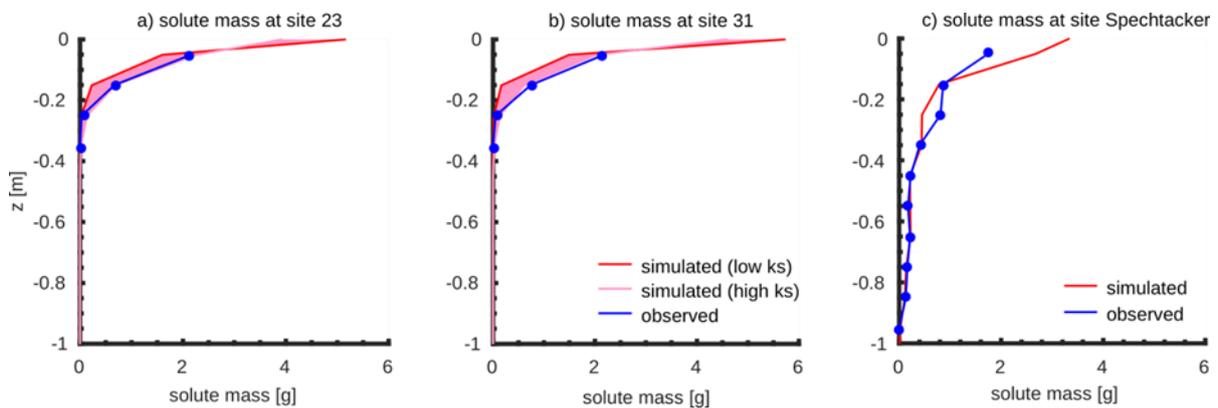


Figure 72: Final Lagrangian Soil Water and Solute Transport Model (LAST)-simulated and observed bromide mass profiles at two sites with dominant matrix flow (a, b) and a site with strong preferential transport (c), after irrigation with 10 mm/h for two hours. Shaded areas mark the uncertainty range caused by the uncertainty in observed hydraulic conductivity. Source: Sternagel et al., (2019).

The most recent version of LAST allows for simulation of reactive transport (Sternagel et al., 2021). Transformation kinetics are simulated by transferring mass from the parent to the child component in each fluid particle according to the reaction rates, as each particle may carry concentrations of different substances. Respective concentrations are limited by the solubility. A retardation coefficient is not helpful in the particle-based framework, as the solute mass travels with the water particles and thus by default at the same velocity. LAST therefore accounts for a reduced solute mobility through explicit transfer of dissolved mass from the water particles at a given depth to surrounding adsorption soil sites (and vice versa). This may either operate under rate-limited or non-limited conditions. If the maximum concentration of the adsorbed substance is locally reached, the remaining solute will travel in a non-adsorbing manner. A first simulation of plot-scale Isoproturon transport observed at the same site in the Weiherbach catchment yielded promising results compared to simulations with Hydrus 1d (Fig. 73).

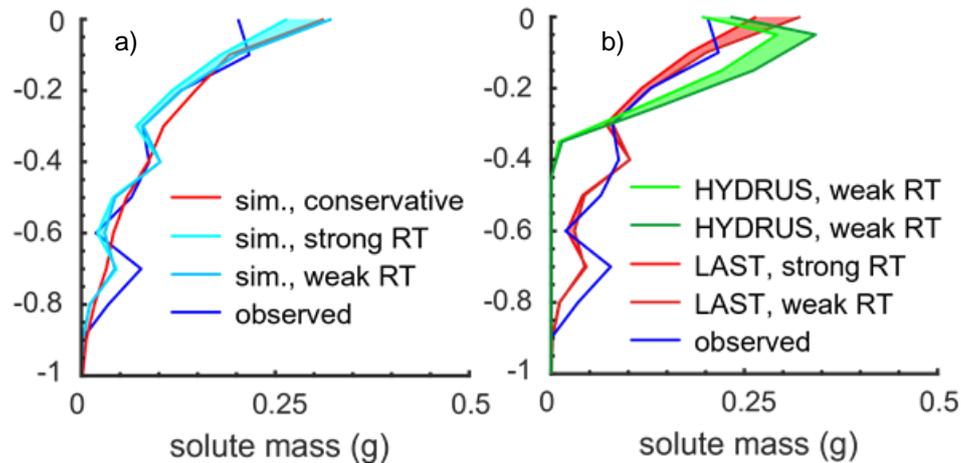


Figure 73: Simulated and observed Isoproturon profiles using LAST at a Weiherbach site dominated by preferential flow (Zehe and Flüher, 2001) 2 days after application (a), revealing differences in simulated conservative and different reactive transport parametrizations due to the variations of the KF and DT 50 values (RT). Benchmark simulations with Hydrus 1d, shaded areas illustrate the variations of the KF and DT 50 values (b). Source: Sternagel et al. (2021).

All these findings corroborate that Lagrangian models provide many assets to simulate flow and transport in heterogeneous soils compared to the traditional Richards and advection-dispersion equations.

#### 2.2.2.3.2 Testing the Stokes flow approach for simulation infiltration across scales and landscape settings

Demand et al., (2019) tested as other alternative to the Richards equation the potential of a Stokes flow infiltration model to simulate sequential infiltration events, using data from the 45 soil moisture sensor clusters discussed in section 2.2.1.3.1. Stokes flow describes 1D vertical flow based on the steady state momentum balance relating the driving gravitational and the hindering viscous forces (Germann and Karlen, 2016). The Stokes flow model relies the thickness of the flow film, controlling gravity, and the specific contact area [ $\text{m}^2 \text{m}^{-3}$ ], which characterises similar to the wetted perimeter in open channel flow friction at the contact line. If the contact area is independent of the rain pulse intensity, the wetting front velocity can be predicted from the input pulse intensity (P). Theoretically, the corresponding relationship of the wetting front velocity and rainfall intensity should follow a power law with an exponent of 2/3 (Hincapié and Germann, 2009). Demand et al. (2019) grouped the observation at the cluster sites according to their geology and landuse into candidate functionally units of similar infiltration, and searched for a typical distinct relationship between rainfall intensity to infiltration front velocity. To assure clearly identifiable rainfall input pulses, they exclusively used for events where the majority of rainfall occurred within 3 h. For these infiltration events, the two parameters of the Stokes flow approach were inversely fitted. Water storage in the soil matrix during Stokes flow was determined from the observed water content waves. Incorporating abstraction of soil matrix into the model is needed to derive water content change, as this is not accounted for in the Stokes flow approach. Using a linear model based on the initial water content and rainfall amount, the wetting of the soil matrix was simulated on at the scales of individual profiles, cluster sites and landscape units.

A total of 1718 infiltration events was simulated using the Stokes flow model. Optimized wetting front velocities differed between different sensor depths, with higher values at greater depths. The wetting

front velocities showed furthermore a dependency on the geology with highest optimized values for Marl and Slate followed by Sandstone.

A comparison to flow velocities inferred from capillary theory revealed, that the observed wetting front velocities were 73 -89% faster. As the capillary forces are much larger than gravity, this finding corroborates that frictional losses during infiltration are much smaller than predicted by soil matrix properties alone. Exponents of the power law relationship between the wetting front velocity and rainfall intensity, ranged between 0.33 and 0.58, depending on depth and landscape unit (Fig. 74).

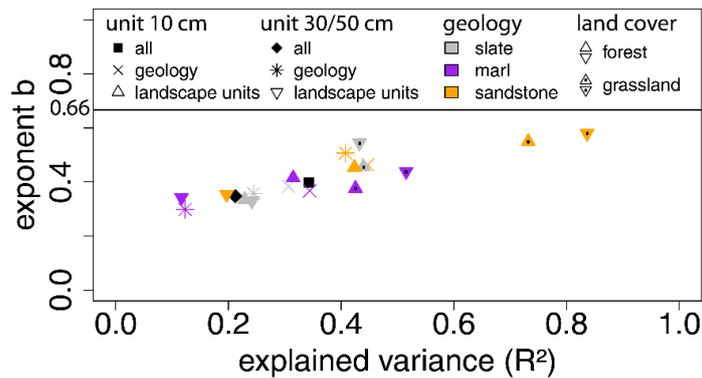


Figure 74: Relationship of explained variance ( $R^2$ ) and the Stokes exponent  $b$  for grouping the sites of different units (all, geology, landscape units) and the two depths. Source Demand et al. (2019).

Especially grassland plots showed exponents close to the theoretical value of  $2/3$  and their fit explained a higher variance. On grasslands in Sandstone, the power law relationship explained up to 84% of the variance, while the explained variance was 42% in forested sites. This was probably due to the higher uncertainty of the rainfall measurements in forests compared to grasslands. A further analysis of velocity-rainfall intensity relationship for candidate functional units (all data, the different geologies, landscape units) revealed improvements in the explained variance for sandstone and also a trend for groups with increasing  $R^2$  to have an exponent closer to the expected  $2/3$ . Overall project D concluded that different dynamic landscape characteristics favor emergence of preferential flow phenomena in the vadose zone. Clayey soils showed the highest susceptibility for preferential flow, which changed with landuse, abundant soil fauna and high rain intensity. This is consistent with the findings of phase II project A (see section 2.2.1.2). Surprisingly, sandstone sites revealed the lowest wetting front velocities.

#### 2.2.2.4 Improving quantitative rainfall estimates and simulations of land surface atmosphere energy exchange fluxes with WRF

##### 2.2.2.4.1 Modell based quantitative rainfall estimates by assimilation of weather radar

Within phase, I project C, an advanced high-resolution modelling and data assimilation system was set up based on WRF-NOAH-MP (Bauer et al., 2015). The model was operated at a horizontal resolution of 3 km in a 2043 x 2076 km large European domain using 57 vertical levels up to 50 hPa. This was beneficial to capture upstream and large-scale conditions with high resolution and consistent representation of physical processes such as convection initiation and cloud-precipitation microphysics. The lower troposphere is represented with 14 levels up to 1500m ASL. Simulations were driven by the hourly ECMWF operational forecasts. Using the WRF 3DVAR data assimilation in a rapid update cycle of 1 hour, a huge data set of conventional observations was assimilated (Fig. 75 a).

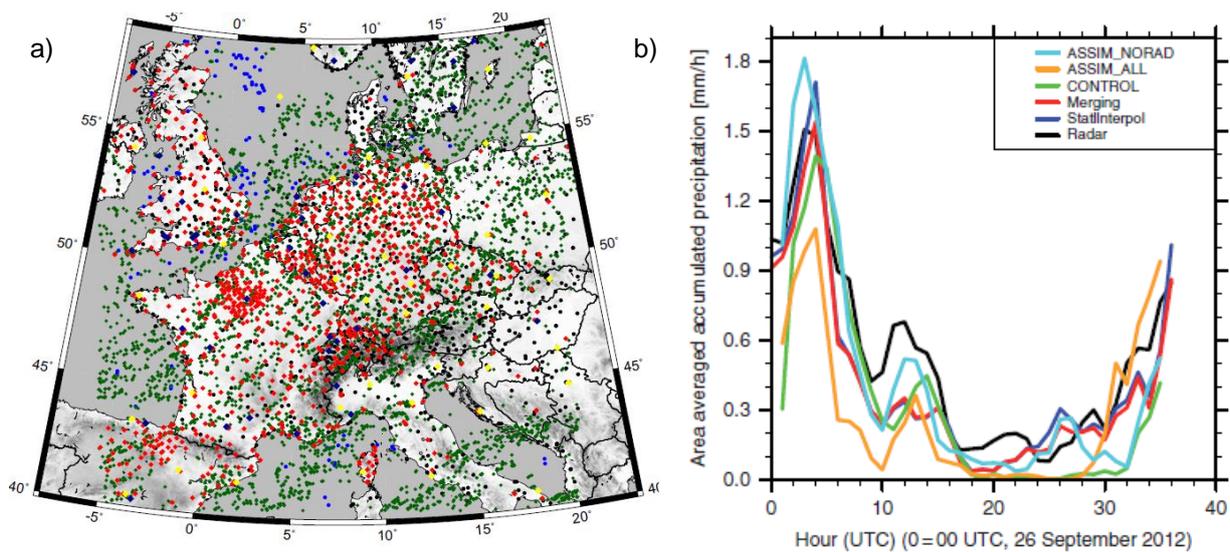


Figure 75: Model domain and data used for assimilation (a): Black = Surface stations (SYNOP + Metar), blue = ship observations (SHIP), green = aircraft observations and atmospheric motion vectors from satellite (AMDAR + SATOB), red = GPS zenith path delay, yellow = radiosondes (TEMP) and brown = wind profiler. (b) Time series of hourly accumulated and area averaged precipitation amounts (mm/h) from 00 UTC, 26 September 2012 until 12 UTC, 27 September 2012 for spatially interpolated rain gauge data (blue), precipitation derived from Wideumont radar data (black), a merged product of the two (red) and three model simulations CONTROL (green), ASSIM\_NORAD (cyan) and ASSIM\_ALL (orange). Source: Bauer et al., (2015).

These include data from weather stations, radio soundings, GPS zenith total delay water vapour estimates, and satellite derived wind fields (AMV). Particularly, radar reflectivity and radial propagation velocity from all operational French radars were used during an important weather situation (26-27 Sep. 2012), where a shift from large-scale frontal rain on Sep 26 towards more convective rainfall at Sep 27 occurred. This made up an interesting test case for comparing the WRF based and geo-statistical QPEs (see section 2.2.1.4). To assimilate radar radial velocities in a variational system, the three model wind components have to be projected to the radar beam position. This was done according to the procedure of Xiao and Sun (2007) by projecting the model wind vector to the line-of-sight of the radar beam. For radar reflectivity, the forward operator of Sun and Crook (1997) was applied, calculating reflectivity from rainwater mixing ratio. Within the assimilation process, the 3DVAR system used a simplified cloud microphysics scheme (Dudhia, 1989). Alternatively, a new forward operator for polarization radar was developed and implemented in the WRF model. This was expected to remove inconsistencies between the two microphysics-schemes in the 3DVAR system and the hourly forecasts and to use significantly more information with respect to the hydrometeor types and their size distributions. After the implementation, the development was tested for a supercell case that occurred in the afternoon of the 30th of June 2012 to the north of Stuttgart.

Although originally intended in the proposal, phase II projects C and E did not continue model-based QPE. This was because results of WRF based QPE performed slightly worse than the QPE that relied on the geo-statistical merging or radar and station data (Fig. 75 b). Given this fact, required substantial WRF modelling efforts were not justified. Instead, phase II project C (group Wulfmeyer) expanded the surprising finding that WRF can be operated at the scales of Large Eddy Simulations (111 m) to explore soil-vegetation-atmosphere feedbacks.

#### 2.2.2.4.2 Coupled simulations of land surface atmosphere energy exchange fluxes with WRF-NOAH-MP-HYDRO

The overarching objective of phase II project C was to better characterize and understand land surface-atmosphere exchange and feedbacks using three different and complementary approaches (compare sections 2.2.1.5 and 2.2.2.1). One of those was to use high-resolution hydro-meteorological simulations of the WRF-NOAHMP-HYDRO model to simulate soil-vegetation-atmosphere (SVA) interactions and their effects on regional evapotranspiration explicitly and test those against the data collected during the field campaign at the study site of Petit-Nobressart (see also Fig. 39). To explore the feasibility of different parametrizations for high resolution modelling, a nested three-domain WRF setup was used. While the largest domain covered the whole of Europe at horizontal resolution of 2.5 km and a vertical discretization of 65 layers, the third (D3) was centred over the Attert catchment at horizontal resolutions of 500 m and 100 m to resolve large eddy turbulence. The boundary conditions for the nested domain were interpolated from the simulation fields of the superordinate domain.

The WRF-NOAHMP model was applied for the day of 2 July 2015 using two different canopy-resistance models i.e., the Ball-Berry scheme or the Jarvis-Stewart scheme and simulations were compared to observations. This revealed that the latter clearly outperformed the former, which yielded a Bowen ratio that was totally in opposite to the observed one (Fig. 76).

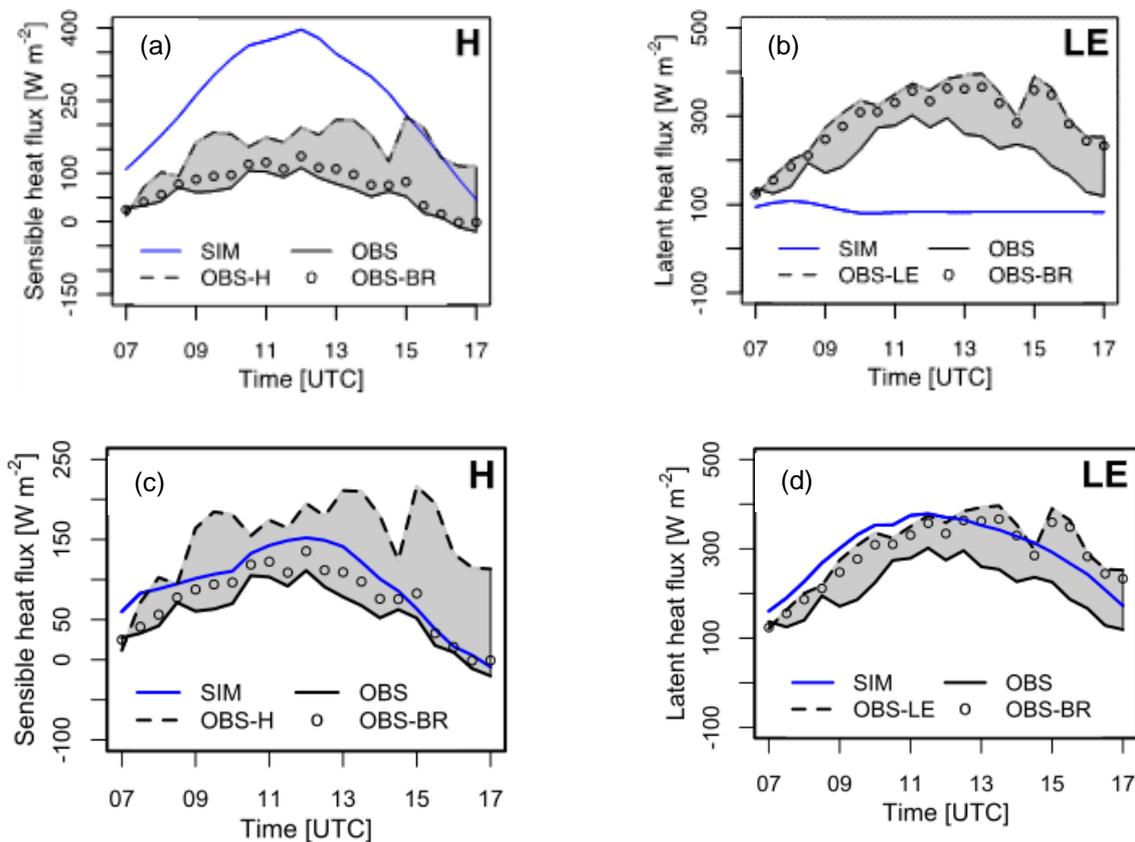


Figure 76: Sensible heat flux  $H$  and latent heat flux  $LE$  on 2 July 2015 between 7 and 17 UTC with the Ball-Berry canopy-resistance scheme (a and b) and the Jarvis scheme (c and d). The shade areas in illustrate the difference between the uncorrected EC data (black solid lines) and the data (black dash lines) corrected by adding the entire residual term to either the sensible (OBS-H) or latent heat flux (OBS-LE). The black circles indicate the corresponding fluxes corrected with the Bowen-ratio method (OBS-BR), respectively.

In order to investigate the reason why the latent heat flux in the Ball-Berry simulation case had a strong daytime reduction, additional simulations with variations of soil moisture and soil temperature were conducted. The reduction of simulated latent heat flux originated from simulated transpiration, which is calculated according to a photosynthesis rate. This reduction of simulated transpiration is due to the temperature inhibition effect, which is hard-coded in the Ball-Berry scheme. This finding is in line with study of Farquhar et al. (1980). The corresponding differences in the simulated sensible and latent heat budgets had a strong impact on the entire structure, wind and moisture dynamics of the boundary layer. For instance, the boundary layer height increases much more rapidly for the Ball-Berry run compared to the Jarvis run. These differences reach values up to 1.2 km in the late afternoon. The upward wind velocities in the Jarvis run are stronger from 10:00 to 12:00 UTC in the lower part of the boundary layer (around 0.5 km) and weaker from 15:00 to 17:00 UTC at higher levels (around 1.5 km) compared with the Ball-Berry run (Fig. 77 a and b). The atmosphere within the PBL in the Jarvis run was dryer in the morning and until noontime and moister afterward, while the atmosphere within the Ball-Berry run dried faster in the morning and kept drying in the afternoon (Fig. 77 c and d).

Overall the WRF-NOAHMP model performed with respect to simulate land surface atmosphere feedbacks much better with the Jarvis canopy-resistance scheme than with the Ball-Berry scheme as further detailed in the final report of phase II project C.

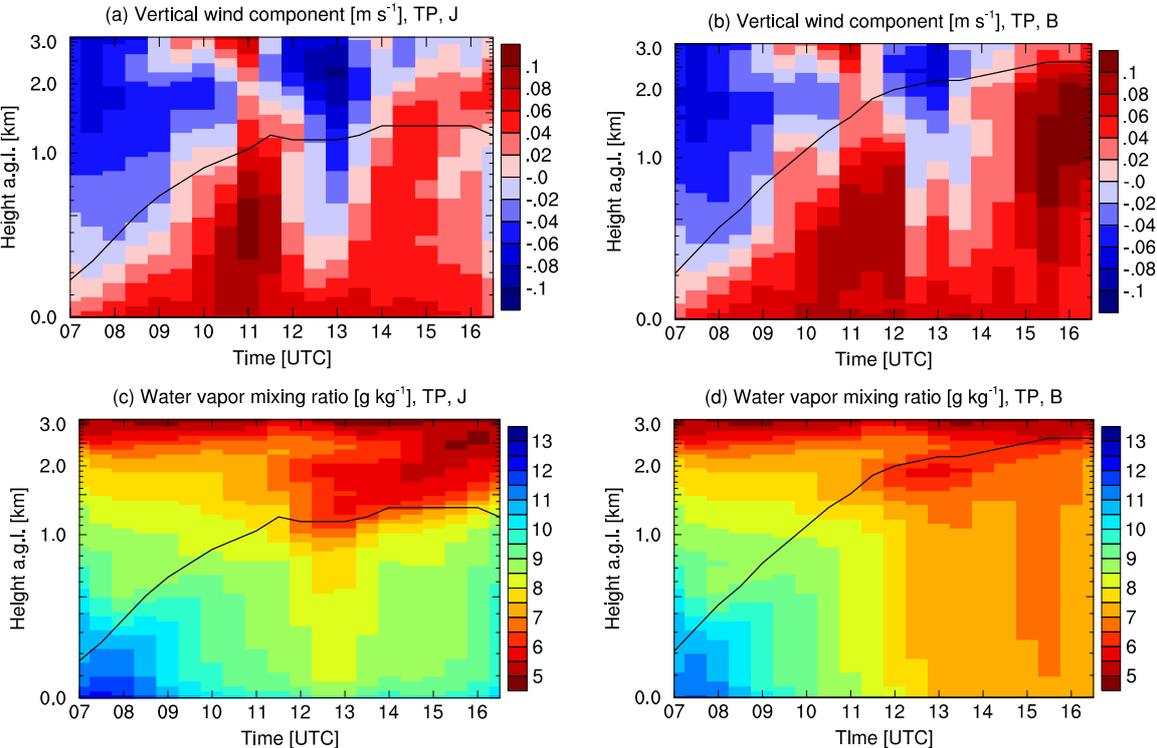


Figure 77: Temporal change of vertical profiles of (a, b) vertical velocity  $w$  ( $m s^{-1}$ ) and (c, d) water vapour mixing ratio ( $g kg^{-1}$ ) from the TP runs with the Jarvis scheme (left column) and the Ball-Berry scheme (right column). The variables are averaged over the Attert catchment for each model layer and averaged for every 30 minutes from seven to 17 UTC on 2 July 2015. The black line in each figure indicates the corresponding PBL height, which is calculated using the bulk Richardson number method with a critical value of 0.25.

### 2.2.2.5 A thermodynamic foundation of the hydrological cycle and hydrological modelling

#### 2.2.2.5.1 Explaining and predicting the surface energy balance, evaporation and atmospheric water cycling in a thermodynamic framework

Research in phase I project D and phase II project C by the overarching question whether thermodynamics adds another constraint to determine land surface functioning, particularly regarding estimates of evapotranspiration, and how such a thermodynamic approach would inform us to better evaluate and understand vegetation optimality approaches. This was motivated by the insight that the turbulent exchange of heat, water vapour and carbon dioxide between the surface and the atmosphere requires physical power to accelerate air in the vertical direction to maintain this exchange. The generation of this power is inherently constrained by thermodynamic principles. However; thermodynamics is practically absent in describing and understanding land-atmosphere exchange processes, which typically use semi-empirical and highly parameterized approaches to describe turbulent fluxes, evapotranspiration and vegetation functioning. An additional, thermodynamic constraint has the potential to radically simplify the formulation of these exchange processes, at least to first order, and in doing so, would allow for a novel, more physically based interpretation of land-atmosphere interactions, vegetation optimality, and their role in hydrologic fluxes. To explore these potentials, project D (and later C) established a thermodynamic foundation of surface-atmosphere exchange and related limitations tested whether those allow feasible predictions of the land surface energy balance across different climate and landuse settings.

The starting point of this theory is the notion that motion requires power to be maintained against friction (Kleidon and Renner, 2013a). This power is derived from differential heating and associated temperature differences. For land-atmosphere exchange, the dominant driver is the absorption of solar radiation at the surface, and the temperature difference is spanned by the surface temperature, at which heat is added to the system, and the radiative temperature of the atmospheric column, which represents the coldest temperature at which the heat can be emitted to space (Fig. 78).

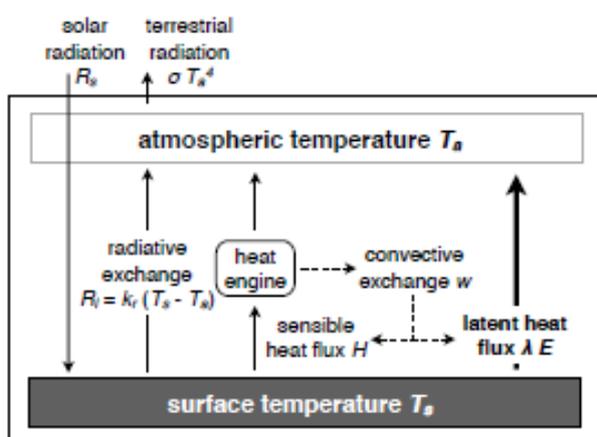


Figure 78: Schematic illustration of the simple energy balance model. Source Kleidon and Renner (2013a).

The resulting turbulent heat fluxes drive an atmospheric heat engine that provides the power to sustain convective motion in the lower atmosphere and the turbulent heat fluxes at the land surface. The power output of this heat engine is constrained by the well-established thermodynamic Carnot limit, which follows directly from the first and second laws of

thermodynamics. This limit is composed of the product of the sensible heat flux and the efficiency of the engine, which is mostly shaped by the normalized driving temperature difference. Applied to land-atmosphere exchange, a critical component is that the efficiency term is not independent of the heat flux as more turbulent exchange implies a stronger cooling of the surface. This reduces the temperature

difference between the surface and the atmosphere, thereby reducing the efficiency of the heat engine and the power output. The consequence is the existence of a maximum in power, with an intermediate value of the heat flux of about half of the absorbed solar radiation, and an intermediate temperature difference between the surface and the radiative temperature of the atmosphere. With a few approximations, this maximum in power can be derived analytically, using absorbed solar radiation as the primary forcing of the surface yet no empirical parameters for the estimation of the turbulent heat fluxes or their partitioning into sensible and latent heat (in the case where the access to soil water does not limit evaporation).

Rather than testing the approach in the first phase in the Attert catchment, where suitable observations were not yet available at the time, it was benchmarked against large-scale datasets to evaluate the approach at the continental scale, derived the surface energy balance partitioning in the climatological mean (Kleidon et al., 2014; Dhara et al., 2016), and quantified sensitivities to global change (Kleidon and Renner, 2013b; Kleidon et al. 2015). The focus on the continental scale and sensitivities to global change provided us with the opportunity to test that the thermodynamic approach works very well across a wide range of climates as well as their response to global change (Fig. 79). The evaporation rate derived from the thermodynamic optimality approach was then used as the conceptual basis for a joint CAOS synthesis for sap flow and soil moisture observations of the first phase of CAOS (Renner et al., 2016) as outlined in final report of phase I project D and phase II project C (group Kleidon).

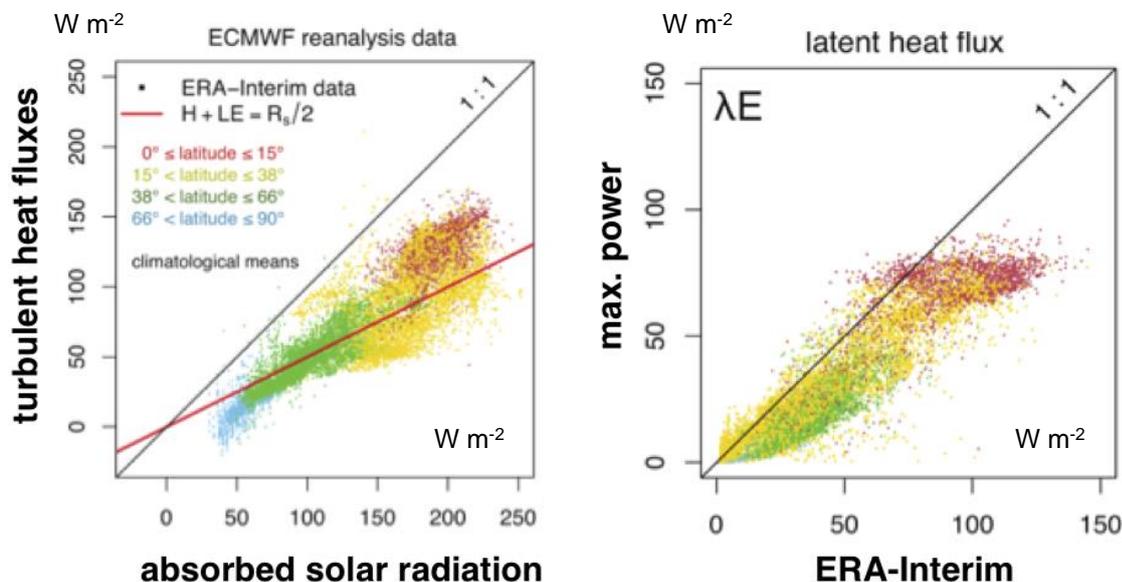


Figure 79: Comparison of annual mean values of (a) turbulent heat fluxes ( $H_{opt} + \lambda E_{opt}$ ) and ERA-interim values to observed absorbed solar radiation and (b) the optimised latent heat flux against corresponding Era-interim re-analysis values. The colour code represents the latitude, lat, of the model grid cell (red: tropics, lat < 15°, yellow: subtropics, 15° < lat < 38°, green: temperate, 38° < lat < 66°, blue polar). Source Kleidon et al. (2014).

During the second phase, the thermodynamic approach was expanded to the diurnal scale to better test its applicability, yet it did not progress to the level to understand effects of heterogeneity and mesoscale circulations as proposed for the second phase. A key conceptual challenge and realization was that at the diurnal scale, heat storage variations that buffer the energy input by solar radiation during the day are primarily buffered within the lower atmosphere in form of the convective boundary layer, rather than

by changes in the ground heat storage. These storage changes are, nevertheless, well constrained by the overall energy balance of the land-atmosphere system, because on average, absorbed solar radiation during the day roughly balances the emission of radiation from the atmosphere to space over day and night (Kleidon and Renner, 2017). Using radiosonde observations from the Lindenberg site in Brandenburg it was shown that this constraint works quite well and that indeed most of the heat storage changes take place in the lower atmosphere (Kleidon and Renner, 2018). The other relevance of this diurnal heat storage variation is that it alters the entropy budget of the atmosphere, resulting in a different thermodynamic limit. This limit looks like a Carnot limit, yet the heat flux that drives the engine is reduced by the heat storage changes that take place within the engine, which is why we refer to it as the Carnot limit of a cold heat engine (Kleidon and Renner, 2018). When this altered thermodynamic limit is used in the maximization of power, the observed diurnal variation of turbulent fluxes across a range of sites, from arctic tundra to tropical rainforest, could be very well reproduced, again, without the need for any empirical turbulence parameters (Fig. 80).

Furthermore, it provided a simple, yet physical explanation why land surfaces warm more strongly with global warming than ocean surfaces (Kleidon and Renner, 2017), a long-standing finding in climate model simulations. Project C attributed this finding to the difference in how the diurnal variation in solar radiation is being buffered over land compared to oceans.

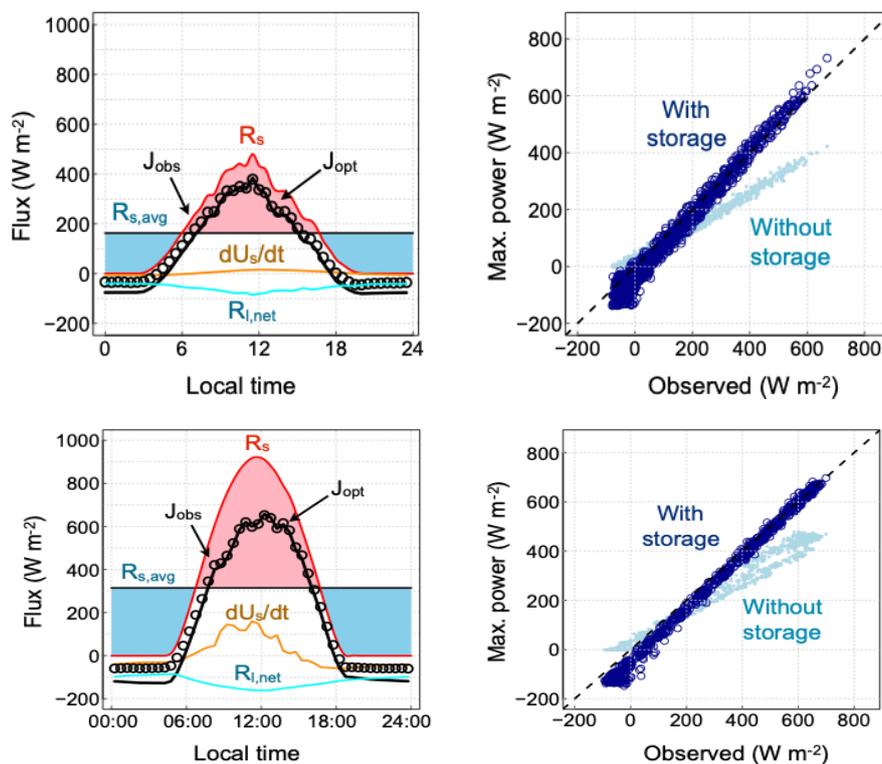


Figure 80: Mean diurnal cycle of the absorption of solar radiation at the surface ( $R_s$ , red line, observed), ground heat flux ( $dU_s/dt$ , orange line, observed), and turbulent heat fluxes estimated by maximum power ( $J_{opt}$ , black line, estimated) and observations ( $J_{obs}$ , black circles, observed) for a selected month for (top) a grassland in a temperate environment in Germany, and (bottom) a planted pine forest in an arid environment in Israel. The comparison of the turbulent heat fluxes estimated from maximum power to energy balance measurements is shown for 30 min observations in the right panel for each site for two cases of thermodynamic limits that differ by their consideration of heat storage changes (dark blue: with storage; light blue: without storage). Source: Kleidon and Renner (2018).

As the buffering of solar radiation takes place above the land surface, turbulent fluxes have a pronounced diurnal variation (as opposed to open water surfaces) and are practically absent at night. Night-time surface temperatures are thus more sensitive to changes in greenhouse forcing, warming more than daytime temperatures. We showed that with the maximum power limit the difference of warming rates should roughly be a factor of two. This interpretation can then not only explain why observed night-time temperatures have increased about twice as strongly as daytime temperatures over recent decades of global warming, but that the combined effect results in land to warm about 50% more than the rate that oceans heat up.

We also tested the sensitivity of our approach to land cover change with observations from Amazonia for tropical rainforest as well as a cropland/bare ground site (Conte et al., 2019). The approach estimated turbulent fluxes and evapotranspiration very well, showing that the tropical rainforest transpired basically at an unstressed rate equivalent to the evaporation rate inferred from thermodynamics during the wet and dry seasons (Fig. 81). The cropland site during the dry season showed a notably lower rate of evaporation due to the lack of vegetative cover, yet the magnitude of turbulent heat fluxes was nevertheless adequately predicted from thermodynamics (Fig. 81).

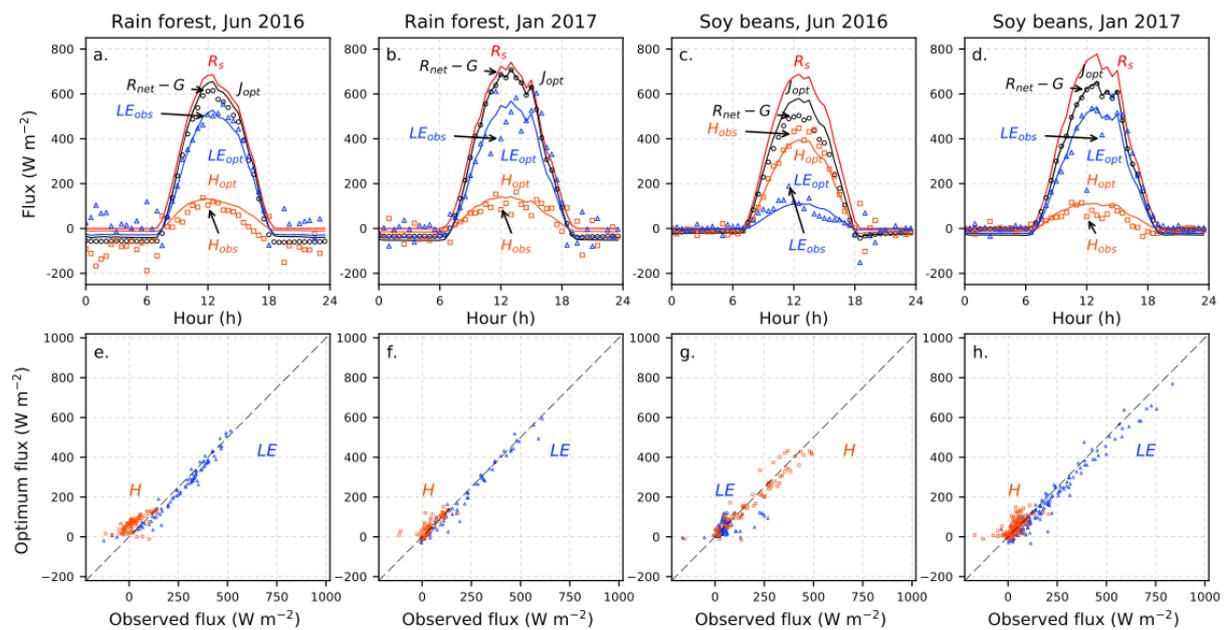


Figure 81: Estimated and observed surface energy balance partitioning for the rainforest and soybean sites for 1 month in the dry and wet season. The top row (a–d) shows the observations of the absorbed solar radiation at the surface ( $R_s$ , red), turbulent heat fluxes (observations,  $R_{net} - G$ , black circles; estimated,  $J_{opt}$ , black line), the sensible (observations,  $H_{obs}$ , light red squares; estimated,  $H_{opt}$ , light red line), and latent heat flux (observations,  $LE_{obs}$ , blue triangles; estimated,  $LE_{opt}$ , blue line). The bottom row (e–h) shows scatter plots between observed and estimated fluxes of sensible ( $H$ , red) and latent heat ( $LE$ , blue). Nighttime values are not shown. Source: Conte et al. (2019).

This outcome is not quite what one would expect from the stark difference in aerodynamic roughness between tropical rainforest and bare ground. Yet, it is consistent with the notion that land-atmosphere exchange is primarily driven by surface heating and buoyant exchange, for which the aerodynamic roughness does not play a dominant role. The thermodynamic approach provides a different perspective on vegetation optimality: The rainforest transpired at the rate predicted from thermodynamics, so stomatal conductance and other vegetation parameters seemed to not have played a role. The two

aspects that were heavily influenced by the rainforest are the absorption of solar radiation (its lower albedo resulted in more of the incoming solar radiation being absorbed), and the sufficiently deep root system made soil water accessible for transpiration during the dry season. It would thus seem that the thermodynamic constraints play a dominant role in limiting vegetation, yet that some factors that affect the thermodynamic constraints, particularly the absorptivity of solar radiation and access to soil water, are modulated by vegetation.

The work on the diurnal cycle and the heat buffering in the atmospheric boundary layer set the conceptual basis for another synthesis effort in CAOS phase II (Renner et al., 2019), where the phase lag between solar radiation and simulated evaporation was used to discriminate the moist suited evaporation model as detailed in section 2.2.1.5.3.

Overall, these findings revealed that thermodynamics provides an additional, relevant constraint to infer energy balance partitioning and evaporation. This essentially requires a systems perspective because interactions with surface temperature set the associated thermodynamic limit, this approach reproduces large-scale patterns of energy balance partitioning and the hydrologic budget as well as the diurnal cycle, and it adequately captures sensitivities to global change.

#### 2.2.2.5.2 Explaining and modelling structural controls on rainfall runoff, preferential flow and storage dynamics in a thermodynamic framework

Within the first funding phase we explored to which extent a thermodynamic perspective may advance and deepen hydrological theory and modelling. Specifically we:

- Tested alternative ways for coupling different process domains and for the executing sequence in hydrological models, by focusing on energy conversion rates associated with hydrological processes.
- Explored whether thermodynamic optimality offers ways for improving model predictions by constraining model parameter sets or even allowing for a-priori uncalibrated predictions.

The first bullet was addressed in phase I project I and phase II project F, which developed and tested a particle based approach for simulating soil water flow in the matrix and in vertical preferential pathways during rainfall driven conditions (compare section 2.2.2.3.1). Mass and energy exchange between both domains was successfully characterized with a thermodynamic re-interpretation of the Bernoulli equation (Jackisch and Zehe, 2018) as further detailed in the corresponding final report.

Westhoff and Zehe (2013) addressed the second bullet in a Monte Carlo study to explore whether thermodynamic optimality, in fact Maximum Entropy Production MEP, provides a feasible constraint for conceptual model parameters of an HBV type model. They found a small intersect between parameter sets matching the water balance and those that maximized entropy production when varying a single parameter of the beta store soil moisture accounting scheme. Unfortunately, the intersection was zero when additionally parameters of the sequence of reservoirs controlling runoff concentration were varied. The reason for this is that water fluxes in the HBV-type model are not driven by potential gradients and our efforts to define suitable proxies were not successful. We concluded that MEP is of low use for reducing parameter uncertainty of conceptual models, because these models are not

thermodynamically consistent. This is a valuable insight to define the necessary requirements for testing thermodynamic optimality approaches with hydrological models.

Westhoff et al. (2014) investigated furthermore the effect of the frequently made assumption of a steady state forcing when searching for thermodynamic optimal model parameters. Starting point was a study of Kleidon and Schymanski (2008), which showed with a simple “electric-circuit-like” model representing the partitioning of rainfall into evaporation and runoff, that one resistance could be inferred using MEP when the other one was known. We showed analytically that the previous assumption of a constant forcing leads to an optimum flow resistance that only depends on the known flow resistance of the other flux, but neither on the annual precipitation input nor on gradients driving runoff and evaporation. As this was not in accordance with our perception, we translated the electrical circuit into an equivalent reservoir model where evaporation is driven by a steeper potential gradient than runoff. When driving this model with a periodic stationary climate forcing to represent wet and dry seasons a second maximum in entropy production emerged; and the latter became dominant in case of long dry seasons. Further analysis showed that the optimum flow resistance and thus the partitioning rainfall into runoff and evaporation depended on the ratio of the gradient driving evaporation and the total rainfall amount. This ratio corresponds to the dryness index in the Budyko framework and a plot of the evaporative index against this ratio resulted in a curve that was highly similar to the Budyko curve (Budyko, 1961). Using an expanded model, Westhoff et al. (2016) showed that the Budyko curve can be derived using the Maximum Power principle as a constraint. We showed that this approach explains offsets of selected water and energy limited MOPEX catchments from the theoretical curve, as optimum curve is sensitive to the duration of the dry season and the amount of months without evaporation (Fig. 82). While this is in line with the findings reported in the previous section, this suggests that actually a family of Budyko curves might exist, which can be inferred from the maximum power principle.

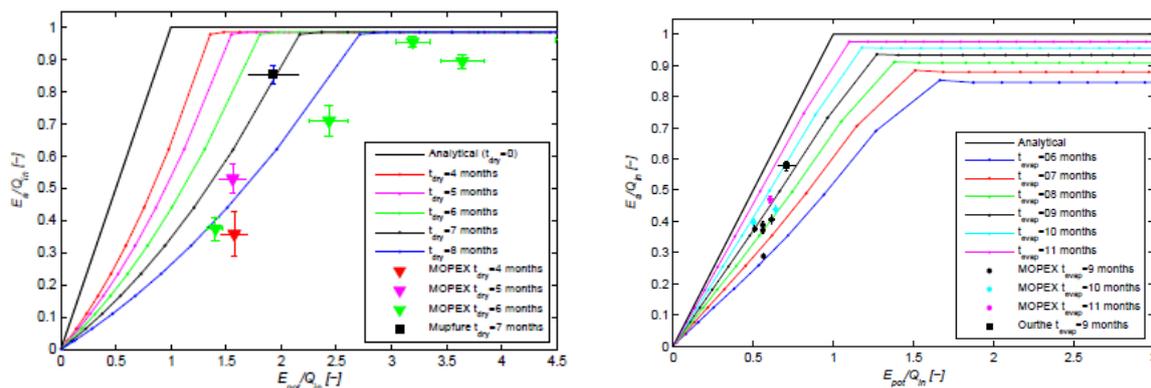


Figure 82: Sensitivity to periodic dry spells in the optimized forward Budyko curve and selected arid MOPEX catchments having at least 1 month with a median rainfall  $< 2.5\text{mm month}^{-1}$  and a coefficient of variance  $< 0.5$  for all months with a median rainfall  $> 25\text{mm month}^{-1}$  (a). Sensitivity to “on and offset in actual evaporation” in the forward model and energy limited MOPEX having a coefficient of variance  $< 0.12$  for monthly median rainfall and with at least 1 month with a median maximum air temperature  $< 0^\circ\text{C}$ . A month is considered to have no actual evaporation if the monthly median maximum air temperature  $< 0^\circ\text{C}$ . Error bars indicate one standard deviation are determined with bootstrap sampling. Source Westhoff et al. (2016).

Along a different line we developed a thermodynamic framework to quantify conversion and dissipation of free energy associated with rainfall runoff and soil moisture dynamics (Zehe et al. 2013). In fact both processes imply conversions and dissipation of capillary binding energy (or matric potential energy) geopotential energy and kinetic energy. Although being very small compared to the land-surface-energy balance, these energy fluxes are the key to explain partitioning of rainfall into runoff and storage from an energy centred point of view, as detailed further below; and they provide a common framework for explaining why preferential flow phenomena occur. Preferential flow leads to faster fluid flows, because preferential flow paths reduce dissipative losses due to an increased hydraulic radius in the rill or river network compared to sheet to overland flow across the land surface (Berkowitz and Zehe, 2020). Subsurface preferential flow of water reduces dissipative losses as well, because friction occurs mainly at macropore or fracture walls, while frictional interactions in the matrix occur along the entire inner surface. Reduced dissipation and faster fluid flow imply a more energy efficient throughput of water, mass and chemical species through the entire system (in case of flow paths spanning the entire system). While this increased energy efficiency is, at first sight, a purely diagnostic observation, it explains the directed and self-reinforcing growth phase of rill and river networks. This is because a higher flow against the potential gradient implies higher power, which means that more kinetic energy can be transferred to the sediments (Kleidon et al., 2013). The related higher erosion rates imply an upstream growth of river networks. A faster flow against the driving gradient means that the latter becomes depleted more rapidly, which in turn means that entropy is produced. This slow negative feedback works against structural growth of river network and this implies the existence of a metastable maximum power state (Fig. 83) as shown by Kleidon et al. (2013).

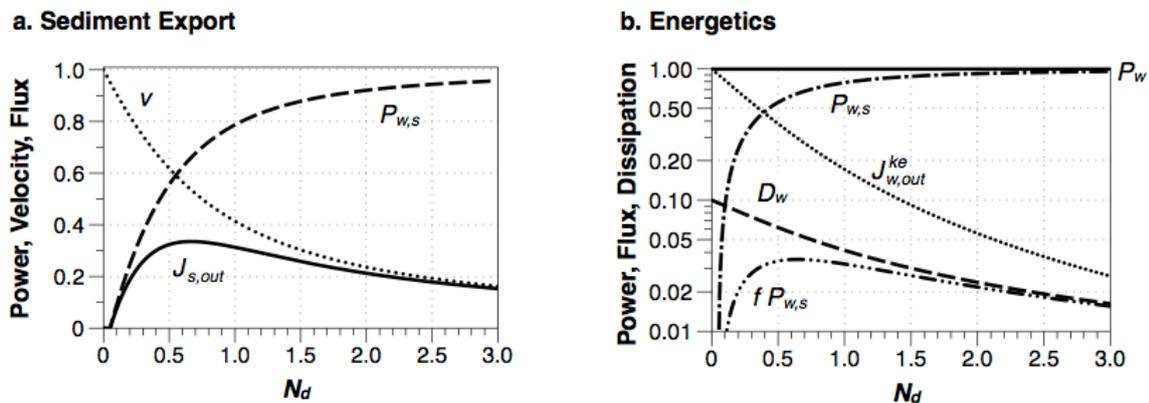


Figure 83: Demonstration of a maximum rate of sediment export resulting from the tradeoff of increased drag resulting in greater work in detaching sediments,  $P_{w,s}$ , but lower flow velocity  $v$ . a: water flow velocity  $v$ , free energy transfer  $P_{w,s}$ , and rate of sediment export  $J_{s,out}$  as a function of the dimensionless number  $N_d$  that characterizes the strength of the drag force,  $F_{w,d}$ , in relation to the accelerating force,  $F_{w,acc}$ , associated with the slope. b: sensitivity of total power  $P_w$ , frictional dissipation  $D_w$  in water flow, kinetic energy export  $J_{w,out}^{ke}$  of water flow, and the free energy transfer  $P_{w,s}$  from water flow to sediment transport, and the fraction  $f P_{w,s}$  that results in sediment export. Source: Kleidon et al. (2013).

In Zehe et al. (2013) we related the total soil water potential to its free energy, and analyzed free energy changes associated with overland flow and soil water dynamics within a physically based model study using the data set of the German Weiherbach. By varying the density of preferential flow paths in soil from zero up to a density where surface runoff completely vanished, we found two local optima that

maximized the energy dissipation averaged across all rainfall runoff events observed within 1.5 y. One of these optima allowed an acceptable prediction of the rainfall runoff response of the Weiherbach without any calibration to discharge data (Fig. 84).

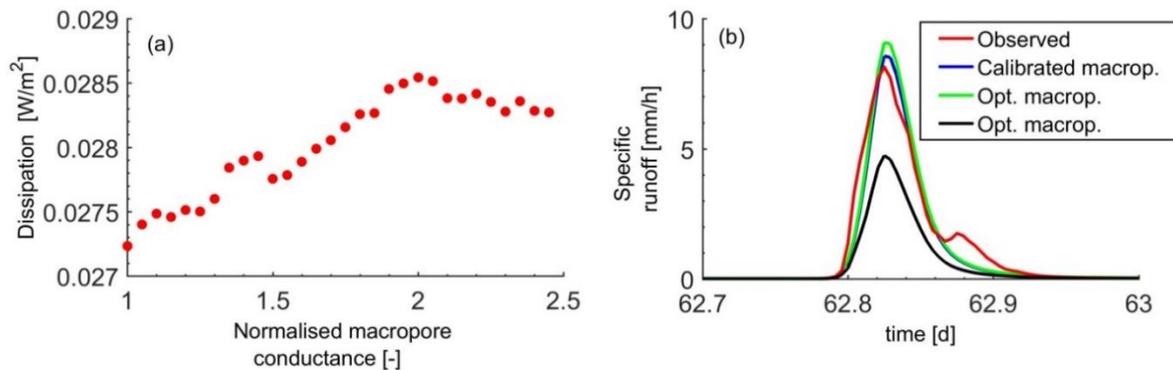


Figure 84: Total dissipation of free energy of soil water as function of the normalized macropore conductance. The latter is total hydraulic conductivity summed of all macropores in the model domain divided to the saturated hydraulic conductivity of the soil matrix. (b) Simulated specific runoff with a macroporosity calibrated to match discharge response of the catchment (Calibrated macrop.), uncalibrated simulations based on the two local optima macroporosities (Opt. macrop.) and observed specific discharge at the catchment outlet. Source: Zehe et al., (2013, adapted).

As one of the optima performed well with respect to predict rainfall runoff, we concluded that the catchment is in an optimal structure as well. The “open point” in the argumentation is that earthworms cannot directly benefit from the enlarged power in soil water fluxes.

Zehe et al. (2013) postulated furthermore the existence of two distinct thermodynamic regimes of rainfall runoff behavior. Cohesive soils work predominantly in the C-regime as free energy dynamics of soil water is dominated by changes in its capillary binding energy; while coarse-grained soils in steep terrain work in the P-regime as potential energy dominates energy conversions in soil. In phase II project F we further explored this idea among others by expanding the thermodynamic framework of rainfall runoff and storage dynamics. The idea is to express the driving matric and gravity potentials by their energetic counterparts and characterize the soil water content by its Gibbs free energy  $e_{free}$ . The latter reflects both the binding state and the amount of water that is stored in a control volume at a given elevation above groundwater and thus relates to the soil retention properties and the topographic setting. Zehe et al. (2019) used free energy to define a new system characteristic determining the possible range of energy states of soil water named energy state function (Fig. 85).

In latter, the soil water storage at local equilibrium, which reflects a state of minimum (zero) free energy, separates the storages corresponding to above mentioned P- and the C-regimes. Wetting of the soil beyond the equilibrium storage implies positive free energy of soil water, corresponding to a state of a storage excess where potential energy dominates. Relaxation back to equilibrium requires that the system releases water e.g. in the form of runoff to deplete the excess in potential energy, and the necessary amount is determined by the overshoot of free energy above zero. Storages smaller than the equilibrium are associated with negative free energy, as the absolute value of the capillary binding energy exceeds potential energy.

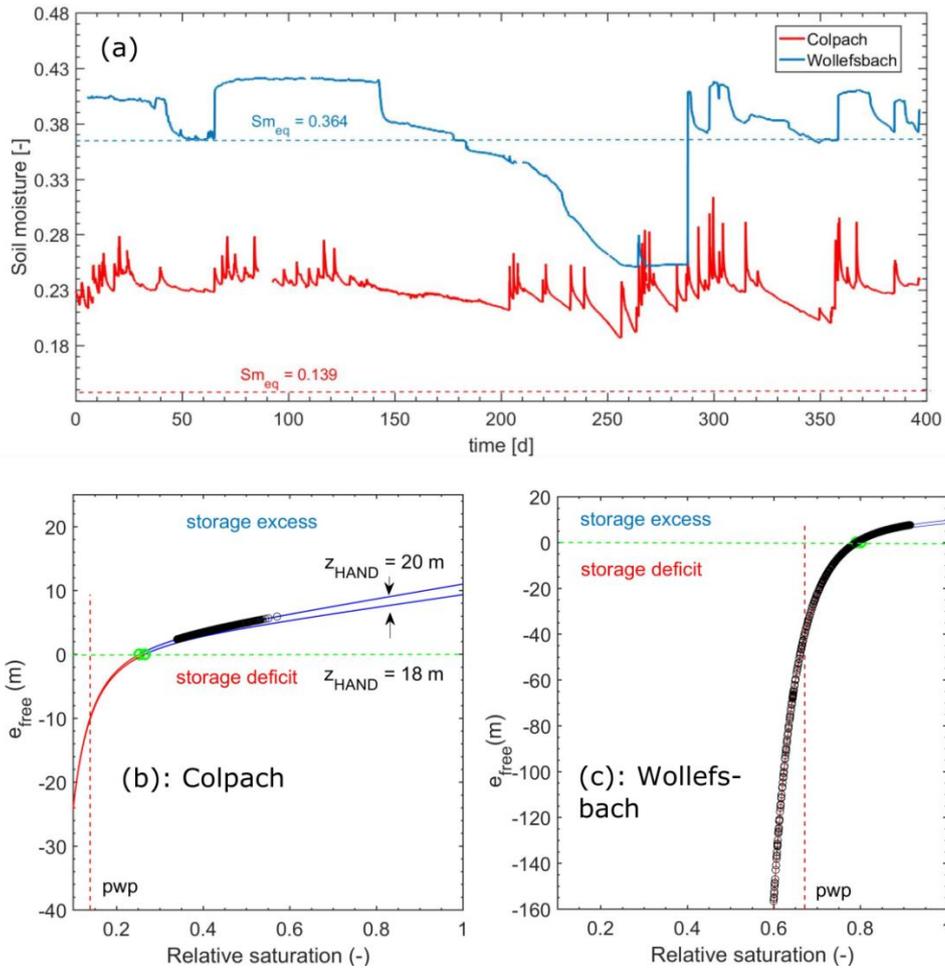


Figure 85: Topsoil water content observed at cluster sites in the Colpach and the Wollefsbach catchment (a) and the corresponding free energy states in their respective energy state curves (b and c note the different scaling of the ordinates). The black circles mark the observations. The vertical dashed line marks the permanent wilting point. Panels b and c show additionally the energy state curve when contamination the real value with an error of minus 2 m ( $z_{HAND} = 18$  m).

Within this C-regime because capillarity acts as the dominant driver for soil water dynamics. The system needs to recharge water to replenish the “energy deficit” below zero, and the necessary amount depends on the distance to equilibrium (compare final report project F). The energy state function defines the possible energy states of the soil water storage, a thermodynamic state space of the root zone. Due to the intermittent rainfall and radiative forcing, their respective annual cycles, the free energy state of soil water will be pushed and pulled through this state space. Zehe et al. (2019) showed that it is straightforward to visualize observed storage dynamics, as pseudo oscillations of the corresponding free energy state in the respective energy state functions. Fig. 85 shows this exemplarily using pairwise soil moisture and matric potential data observed at a cluster site in the Colpach and the Wollefsbach, which are both located 20 m above their respective streams. This corroborates together with Fig. 21 in section 2.2.1.3.1 that storage in the Colpach operates mainly in the P-regime of a storage excess, while clay rich soils in the Wollefsbach are characterized by distinct storage deficit during the summer period.

The free energy state of soil water in the riparian zone of both study catchments has also been proven rather helpful to explain the threshold like onset in streamflow generation (see also section 2.2.1.3.4).

We found a distinct threshold behavior for storage controlled runoff production in both catchments, and clear hints at Hortonian overland flow contributions in the Wollefsbach - the onset of a potential energy excess of soil water in the riparian zone coincides with the onset of storage controlled streamflow generation (Fig. 86). While we admit that a threshold like dependence of the onset of runoff generation is frequently reported, we like to stress that this tipping point has a theoretical basis.

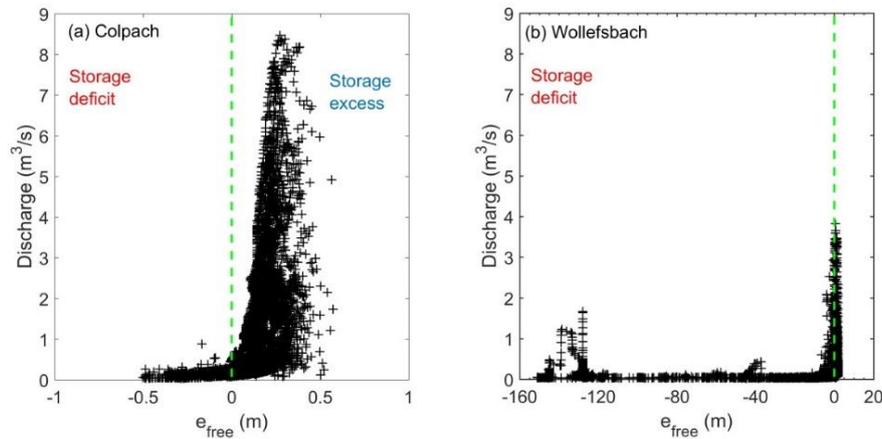


Figure 86: Observed stream flow in the Colpach (a, drainage area is 19.4 km<sup>2</sup>) and the Wollefsbach (b, drainage area is 4.5 km<sup>2</sup>) plotted against the free energy of sites in their corresponding riparian zones.

Complementary to that Loritz et al. (2019) proposed a novel thermodynamic index to better explain distinct differences in runoff generation between the shist/slate and the marl geological settings (Fig 87.)

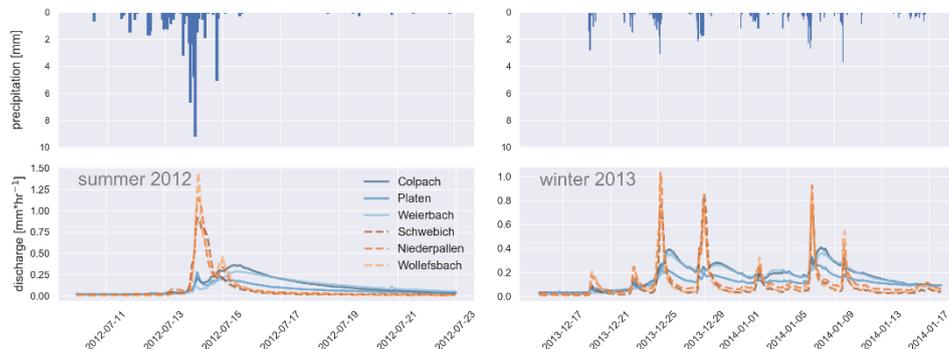


Figure 87: Observed specific discharge and precipitation with different ordinate scales for a period in summer 2012 and winter 2013 (orange: marl catchments and blue schist catchments. Source: Loritz et al., (2019).

Given the fact that potential energy differences are the main drivers for runoff generation, topography provides important information to explain runoff generation in catchments. However, due to the strongly dissipative nature of runoff generation, the driver of a flux explains only one aspect of the runoff generation. In fact, dissipative losses dominate runoff generation and even in case of overland flow only a tiny portion of the driving potential energy is transformed into the kinetic energy of runoff. We thus proposed a new topographic index named reduced dissipation per unit length (rDUNE), which jointly accounts for the energetic driver and the dissipative loss along the flow path. rDUNE allowed a superior discrimination of six catchments into groups of similar runoff generation than the often recommended height over next drainage (HAND) or the famous topographic wetness index (TWI) (Fig. 88).

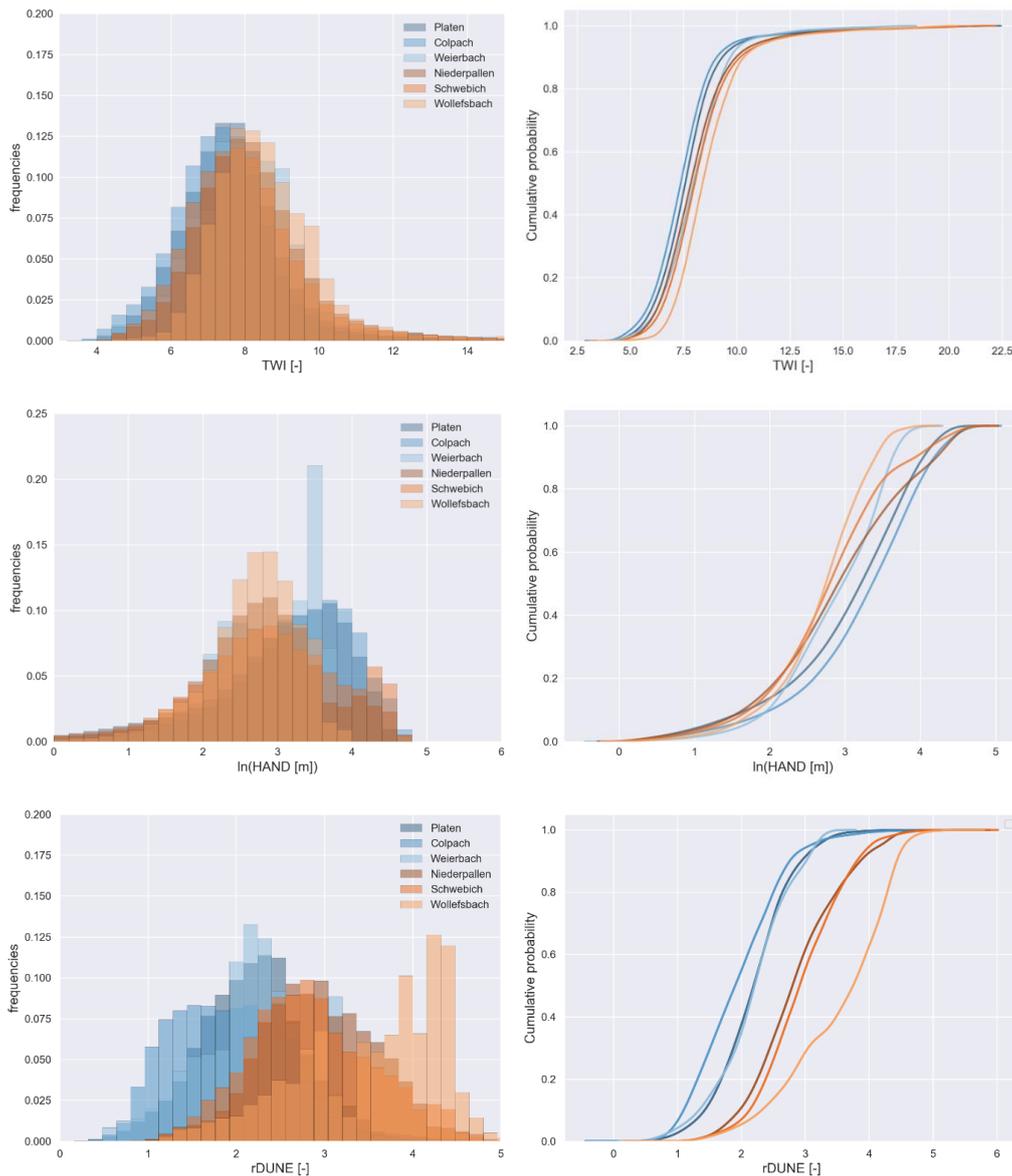


Figure 88: Frequency distributions and cumulative density functions of the TWI, ln(HAND) and rDUNE for the schist catchments in blue the marl catchments in orange. Source: Loritz et al. (2019).

Overall, the reported findings corroborate that a thermodynamic perspective on hydrological systems offer holistic information for explaining soil water dynamics and runoff generation and promising ways for constraining hydrological models that cannot be inferred from the traditional water balance thinking.

#### 2.2.2.6 Information as universal measure to quantify uncertainty, similarity and organisation

Complementary to the thermodynamic avenue, phase II projects E and F explored the benefit of an information perspective along several lines. One of the main challenges of working in an interdisciplinary research group is to handle the large variety of data not only from the point of view of storage and exchange, but also in terms of finding a common framework for evaluating the value of these data. We found information theory, to be extraordinarily well suited for this task. In information theory, the value of data can, irrespective of its units and disciplinary origin, be expressed as information content, which can also be interpreted as uncertainty reduction. Adding to these advantages, (algorithmic) information

theory offers ways to express the information content of both data and models in a single quantity, bit, which is also the fundamental unit when processing data or running models on computers, nowadays the main tool of science. We found information theory applied to the problems of hydrology in general and to those of our research group in particular to be an extremely powerful tool as already shown above. It opened a completely new perspective of looking at data and models, allowed direct comparison of the usefulness of all sorts of data and all sorts of models, from data- to physics-based (Fig. 89), and provides via information entropy a natural link to thermodynamics and organisation.

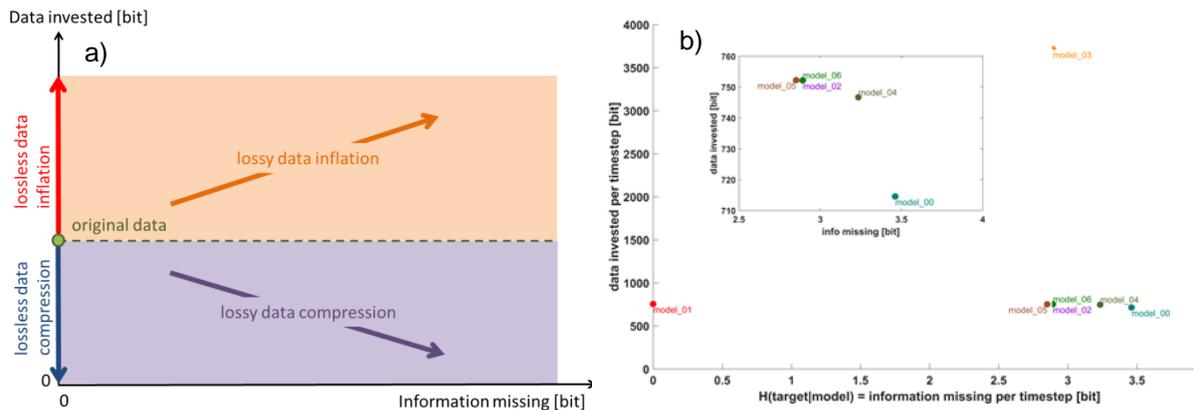


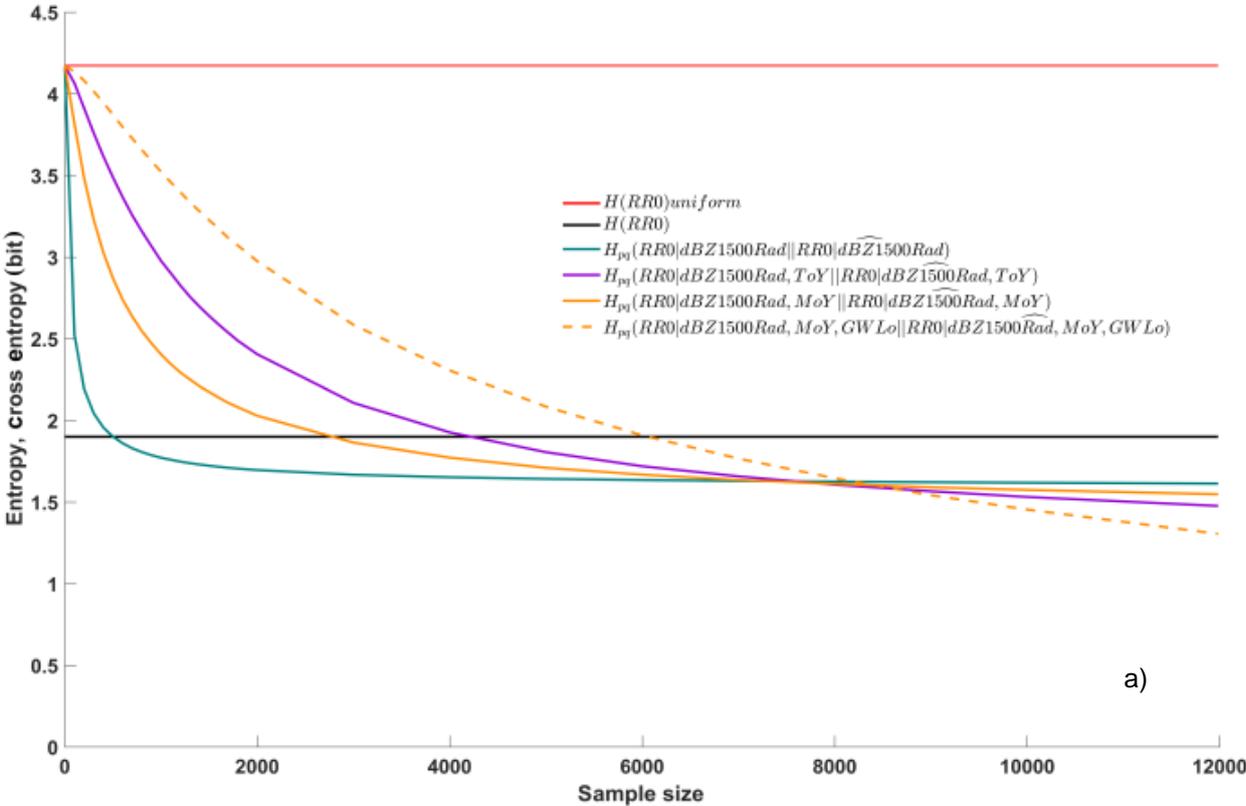
Figure 89: Generalized evaluation of models in terms (a) of efficiency (data invested) and effectiveness (information missing). Model efficiency vs. model effectiveness for six hydrological models differing in terms of structure, resolution, and numerical scheme.

Originally, we planned to use information theoretic measure for the inter comparison of various hydrological models, namely CATFLOW (project F), the CAOS model (project E), ROGER (project D) and NOAH-MP (project C). The idea was to develop a framework to evaluate which of them makes best use of available data, shows the best performance at an appropriate complexity. The CAOS model was conceptualised, code and set up for the entire Attert basin, as detailed in section 2.2.2.1, but so far the simulation quality not sufficient for such an exercise. Furthermore, project D neither succeed in applying the ROGER nor did project C set up the NOAH-MP model for the Attert. The CATFLOW model was successfully setup for the Colpach using 105 hillslopes, and here we used information entropy to shed light on the time varying dynamic similarity of simulated runoff and storage, as well as a new mutual information index as metric for clustering functionally similar hillslopes into functional units (section 2.2.2.1.3). We also designed and coded a new conceptual model called SHM (Simple Hydrological Model) and used this as a testbed for the development of dynamical adaptive model clustering using again information entropy as means to quantify similarity of model elements as detailed in section 2.2.2.2. As our framework for model evaluation turned out to be much more generally applicable than envisaged, we decided to shift from the original set of models to a more diverse range encompassing purely databased and conceptual models. Our main insight was that with just a few well-chosen predictors (mainly precipitation sums and gauge observations of the immediate past), purely databased models perform as good as conceptual models. While this is somewhat disenchanting for conceptual hydrological modellers, it demonstrates how much information about discharge is contained in both the meteorological drivers and in the memory of the system. Concepts from information theory were also found to be very helpful useful for the model based effort of Rodriguez et al. (2020) to infer catchment

scale travel time distribution. The used the Kullback-Leibler-Divergence to quantify the information gain when a posteriori constraining travel time distributions inferred from either deuterium or tritium with the other independent data set (compare Fig. 26, section 2.2.1.2.3).

We used furthermore information concepts to evaluate the homogeneity of clusters of similar soil moisture patterns (Mälicke et al., 2020; see Fig. 20 section 2.2.1.3.1), delineate rainfall runoff events (Thiesen et al., 2019) and for improving the geo-statistical merging of radar and rain gauge data into a high quality precipitation estimate (Neuper and Ehret, 2019, see section 2.2.1.4). Specifically, we compared the value of various observables for estimating ground precipitation (Neuper and Ehret, 2019). Using the language of information theory, we were able to quantify the information content of data and the effect of sample size in the same unit, bit, and evaluated these results using maximum entropy estimates as benchmarks. Fig. 90 corroborates that a careful choice of predictors is essential for robust QPE, as increasing the number of predictors quickly demands very large learning data sets. We found that ground observations proved the most informative predictor for distances up to 8 km, beyond quality-checked weather radar data were superior.

Overall, we conclude that information theory is a superior means to evaluate information content of data, uncertainty and redundancy in multiple simulated and observed time series. Information entropy is furthermore, ideal to quantify space-time organisation in hydrological system properties and dynamics in a very intuitive manner and thus and ideal complement to a thermodynamic perspective.



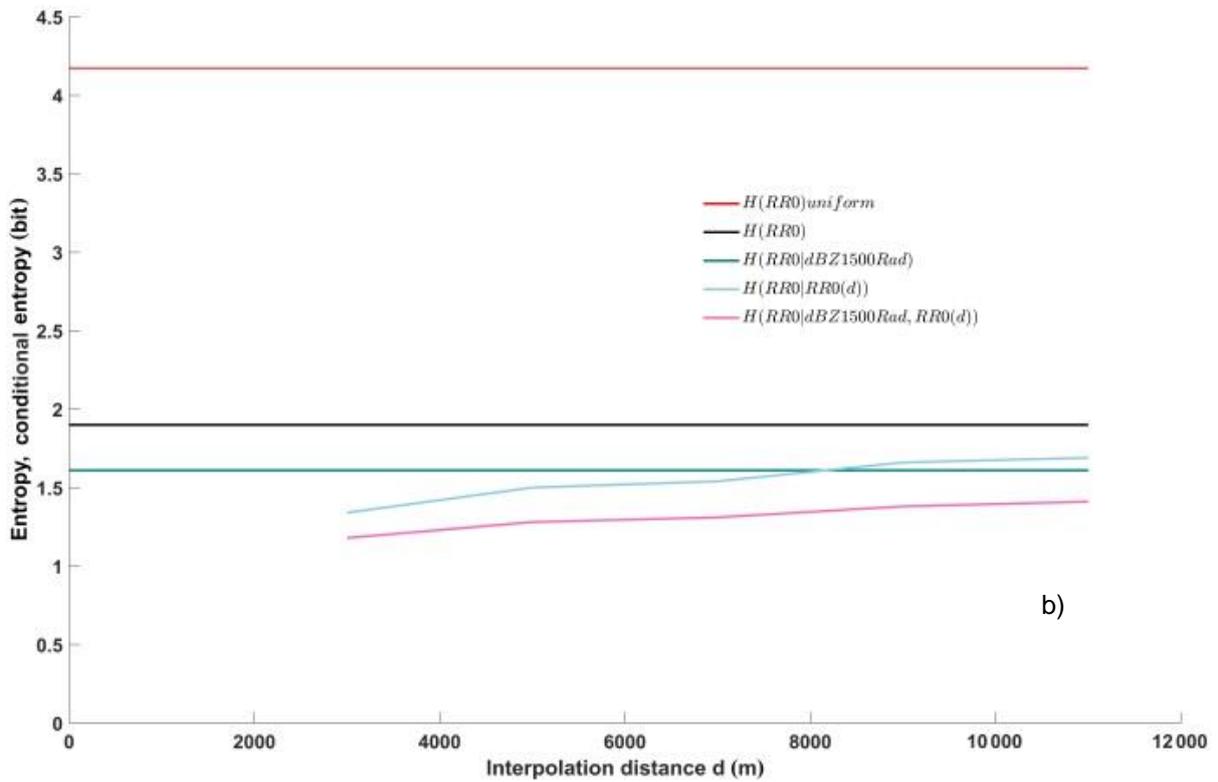


Figure 90: Entropy of ground rainfall for no predictors available (a, black line), entropy of the benchmark maximum entropy uniform distribution (red line), cross entropies between conditional distributions of the target given one, two and three predictors for the full data set and samples thereof (green, purple, yellow and dashed yellow lines, respectively). The “^” symbol indicates a sample. Entropy of ground rainfall for no predictors available (b, black line, same as in left panel); entropy of the benchmark uniform distribution (red line, same as in left panel); conditional entropy of the target given reflectivity as predictor (green line); conditional entropy of the target given station rain rate observations as a function of interpolation distance (light blue line); and conditional entropy of the target given reflectivity and rain rate at stations as a predictor as a function of interpolation distance (pink line).

### 2.3 Central issues and research data management

Joint research and field work was a continuous inspiration to the entire CAOS team. Beside a high number of bilateral meetings and the large joint field campaigns to setup the cluster network, the joint irrigation experiments in phase 1 and 2 and the earth worm survey, the science retreats of the entire CAOS team were highlights to share recent findings and discuss new ideas. In the kick off phase these meetings were rather dense: three per year in 2012 and 2013. After that, we switched to annual meetings.

The CAOS data is for sure one of the most diverse and complete research data sets in the world. A professional research data management was thus from the beginning a key backbone activity of the synthesis project S and later on of the CAOS team. Unfortunately, DFG did not provide any substantial funding for this. Furthermore, when trying the CUASHI or the TERENO data base systems, we found the metadata models way too inflexible for the CAOS dataset. The key to substantial progress in this respect was the successful acquisition of the E-Science project V-FOR-WaTer by the group of the speaker Erwin Zehe. V-FOR-WaTer<sup>1</sup> has been developed for almost five years, starting within the E-Science initiative of the Ministry for Science, Research and Arts Baden-Württemberg (MWK), to foster professional management of hydro-meteorological data and data-based learning in a virtual research environment. V-FOR-WaTer is a joint development of the Steinbuch Centre for Computing (SCC) and the Institute for Water and River Basin Management (IWG) at the Karlsruhe Institute of Technology (KIT). The product of this first development phase is a B-prototype of a web portal, including a digital map, sophisticated data filters, workflows and visualization of time series. Moreover, V-FOR-WaTer offers a set of tools, ranging from simple signatures (flow regimes, flow duration curves), univariate statistics to a SCI-PI package for comprehensive variogram analysis and ordinary kriging (which has 100 users today). The data model allows management of a wide range relevant hydro-meteorological data and with a flexible metadata catalogue, which is specifically well suited for the CAOS dataset. The latter is currently integrated into the system to provide FAIR open access to the CAOS data

### 2.4 Qualification of young researchers in CAOS

PhDs, MSc. and BSc. Theses:

- Phase II Project A and phase I project J: 1 Post Doc (promoted to Assistant Professor today), 1 PhD completed, 2 MSc and 2 Bsc.
- Phase II Project B group Tronicke and phase I project F: 2 PhD completed, 1 Post Doc., 2 MSc. and 2 BSc. theses completed
- Phase II Project C group Kleidon and phase 1 project D: 1 Post Doc, 1 PhD completed, 2 MSc. theses completed
- Phase II Project C group Schulz and phase 1 project B: 3 PhD completed, 3 PhD ongoing, 2 MSc. theses completed
- Phase II Project D and phase I Project H: 3 PhD completed, 2 MSc. theses completed
- Phase II Project E and phase I Project S: 1 Post Doc (promoted to Assistant Professor today),

---

<sup>1</sup> <http://www.vforwater.de/>

3 ongoing PhD theses, 5 MSc. theses completed;

- Phase II Project F and phase I Project I: 1 PhD-Post Doc transition (promoted to Junior Professor), 3 PhD completed, 4 MSc. and 1 BSc. theses completed;
- Phase II Project G and phase I Project G: 1 PhD-PostDoc transition, 3 PhD ongoing, 1 MSc. and 1 BSc. theses completed

International summer and winter schools:

- 3 days winter school 2015 (March) Axel Kleidon, Erwin Zehe (Thermodynamics in Earth system science)
- 4 days summer school 2018 (March) Rui Perdigão (Dynamic System Analysis in Environmental Sciences)
- 2 days summer school 2018 (March) Axel Kleidon, Erwin Zehe (Thermodynamics in Earth and its subsystems)

Outreach: CAOS in schools

- Classes on rainfall formation and characterisation of drop size distribution by Malte Neuper on the primary school level (Project E)

### 3 Summary and conclusions

Joint research within the CAOS project explored the link between spatial organisation and hydrological functioning of lower mesoscale catchments for almost eight years. Hydrological landscapes exhibit considerable spatial organization, which manifests through structured patterns of topography, soil, vegetation and most prominently through ubiquitous surface and subsurface drainage networks (Zehe et al, 2014). While this organised “catchment form” strongly determines present storage, cycling and release of water, energy and nutrients, the catchment form has been shaped by water, energy, and nutrient flows of the past (Dietrich and Perron, 2006; Troch et al, 2013). Our main hypotheses was that this co-evolution of form and functioning is reflected in the existence of hierarchy of functional units, which exhibit a distinct hydrological behaviour with respect 1) to the land surface energy balance (Elementary Functional Units) and 2) hillslope scale runoff generation (**hypothesis H1**). Joint experimental and model research in the Attert experimental observatory explored whether:

- Such a hierarchy of functional units exists and whether their typical structure and dynamic behaviour of functional units can be characterised through replicated experiments and monitoring (**hypothesis H2**).
- These functional units are building blocks of simplified and yet physically consistent hydrological models of adaptive complexity for lower mesoscale catchments (**hypothesis H3**). An adaptive model complexity was deemed to reflect the time variant complexity of hydrological dynamics and to save unnecessary computation time by avoiding redundant computations.

In the second funding period these guiding hypotheses were integrated into a set of four broader research questions and related deliverables, in order better reflect and integrate joint research interests and to foster the most promising research activities of the first funding period:

- Q1: How to define, characterize and link hydrological functioning and underlying controls across a hierarchy of spatial and temporal scales, including their interactions with the soil, vegetation and the atmosphere?
- Q2: Does spatial organization in catchments imply the existence of a hierarchy of functional units, which act similar with respect to different hydrological functions and can we characterize their typical behavior in an exemplary manner?
- Q3: How to balance necessary complexity and falsifiability of catchment models to step beyond the input-output paradigm?
- Q4: What is the connection between thermodynamics, organizing principles, catchment structure and catchment functioning?

In the following we synthesize our main scientific achievements and misses, in relation to the four overarching research questions Q1 – Q4 and the three overarching hypothesis.

#### 3.1 Q1: How to define hydrological functional similarity and its relation to landscape structure? D1: Connect landscape organisation and hydrological functioning.

The combination of hydrological process research with concepts drawing from thermodynamics and information theory turned out to be most helpful for defining functional similarity and linking this to structural properties and spatial organisation of the landscape. Hydrological processes are on a very

general level dissipative, as storage, mixing and release of water as runoff and latent heat are generally associated with conversions, dissipation and export of free energy. The thermodynamic perspective is hence the natural choice to understand the interplay of the mass and free energy balance of hydrological systems, explain their dependence on the prevailing forcing, landscape structure and state. While fluxes of mass, energy, momentum or entropy are driven by potential gradients (e.g. in temperature and hydraulic head), it is the flow resistance in the control volume that determines how much energy the fluid dissipates when flowing through the soil, macropore, river net, tree or aquifer. Key emphasis of our experimental design was hence on the characterisation of those landscape characteristics and state variables, which control gradient and resistances, with a strong emphasis on preferential pathways. Preferential flow paths reduce dissipative losses and thus the control volume resistance and imply a more energy efficient redistribution and release (Berkowitz and Zehe, 2020). The thermodynamic perspective was also most fruitful for model simplification and scaling and testing energetic optimality and limitations as constrains hydrological predictions as further detailed when we reflect on question Q4 in section 3.4.

In phase II, we combined the thermodynamic approach with an information perspective. Information theory is ideal to provide a common framework to evaluate the value of data, because it can, irrespective of its units and disciplinary origin, be expressed as information content, which can also be interpreted as uncertainty reduction. Information theory applied an extremely powerful tool, for instance to identify rainfall runoff events (Thiesen et al., 2019), designing a general framework for model evaluation (project E) and exploring the feasibility of hydrological models with adaptive complexity (Ehret et al., 2020). Information theory is well suited for diagnosing the degree of redundancy a) in a distributed model setup of 19.4 km<sup>2</sup> large Colpach catchment consisting of 105 different hillslopes (Loritz et al., 2018) and b) in soil moisture time series at the cluster sites (Mälicke et al., 2020). In the former study, Shannon entropy was used as measure for diversity and thus uncertainty of the simulated runoff of the hillslope ensemble at each time step. Although the entropy of the ensemble was rather dynamic in time, it never reached the maximum value. The latter implies that hillslopes contribute in a unique fashion to streamflow, while a zero entropy means that all hillslope yield an indistinguishable runoff response. Functional similarity is, as stated in our central hypothesis, nothing static but a dynamic attribute, which jointly reflects changing spatial organization of processes in the landscape and their dissipative nature. Information entropy is hence well suited to diagnose the degree of organization of a system state or a process based on its deviation from the entropy maximum (Berkowitz and Zehe, 2020; Zehe et al., 2021), because its definition is equivalent to Gibb's definition of physical entropy in statistical mechanics (Ben-Naim, 2008). We thus conclude that uncertainty about and organization of a system state are two sides of the same medal related.

With respect to the more traditionally concepts to characterize and define hydrological functioning of catchments we explored the concepts of functional connectivity distributions and transit time distribution, the former experimentally, the latter with multiscale experiments and models. This yielded new promising techniques to measure connectivity as well as clear on the role of shallow ground water on hillslope stream connectivity (Beiter et al., 2020). Yet these findings are largely event based and descriptive and thus not straightforward to generalize. Compared to that, our findings on transit time

distributions appear much more general both with respect to the modelling approach based on StorageSelectionFunctions and the Master equation (Rodriguez et al., 2021) and the often forgotten importance of the vadose zone as key control on overall transit time distribution of stream flow (Sprenger et al., 2016).

### **3.2 Q2: Does spatial organisation in catchments imply the existence of functional units? Functional units/ D2: Experimental protocol to characterise structure and functioning.**

The studies of Loritz et al. (2017, 2018) largely confirmed the existence of functional units of similar runoff generation in line with hypothesis H1. In fact we found that 105 hillslope models yielded highly redundant runoff time series (and even more redundant storage dynamics), which could be grouped into six different functional units using mutual information of the entire discharge time series as objective function. Differences between these functional units were well explained by new index r-Dune. Functional similarity was, as postulated, not static but dynamic as reflected in the entropy of the corresponding discharge simulations (see previous section). In line with this, we showed that the concept of an adaptive model complexity, which avoid redundant computations without performance loss, is feasible for a conceptual model (Ehret et al., 2020) and a distributed physically based model (Loritz et al., 2021). This is very much in line with hypothesis H3. Functional similarity with respect to the susceptibility to vertical preferential flow was also corroborated experimentally, based on the analysis of non-sequential soil moisture responses in the 135 soil moisture profiles (project G). This revealed, however, that the combination of geology (and thus soil type) and landuse was best suited to discriminate functional similarity here. Forests were more susceptible than grassland, in summer non sequential recharge (NSR) events occurred more frequently than in winter, the Schists profiles revealed the largest amount preferential flow events. The latter was corroborated by the corresponding dye and bromide tracer transport profiles (projects A and F). We thus conclude that is co-evolution of catchment form and function caused hillslope indeed scale of functional units (Zehe et al, 2014), which exhibit dynamically similar runoff generation (Loritz et al., 2018). Despite these finding, we must admit, that the existence of Elementary Functional Units of similar land surface – atmosphere exchange and related vertical flow processes as spatial entities could not be confirmed. We conclude that the anticipated extend is likely too small compared to critical length scales in the atmospheric boundary layer, as lateral mixing in atmosphere dominates against variability of land surface atmosphere exchange at the 10m scale.

Preferential flow phenomena were generally a continuous inspiration for developments and testing of alternative theoretical and model concepts: the Lagrangian model approaches LAST and echoRD (project F) or the Stokes flow approach (project D) to better predict this still enigmatic phenomenon. Similarly, we experimentally corroborated that subsurface preferential flow phenomena in downslope direction are of key importance for runoff generation in the Schist geology. In this context the combination of time lapse GPR and TDR proved as very power-full tool to visualise and quantify vertical and lateral subsurface flow (projects B, A, F and G). The combination of piezometers, salt tracers and ERT was also feasible to detect and visualise this phenomenon (projects G and B). One of the most exciting experimental breakthroughs is the new irrigation device, which makes four-dimensional time lapse GPR during rainfall experiments possible. This allows for in-situ monitoring of infiltration and quantification of soil moisture changes from relative reflectivity changes and suitable petrophysical

relations (project B, together with A and F). Unfortunately, we did not live up to the expectations of project B to link remote sensing signatures to GPR and to use the former to regionalise the latter to larger scales. This is largely because the envisaged setup of a cable car in the envisaged focus hillslopes was not feasible.

The sensor clusters provided, as expected, a treasure of diverse hydrological data and an anchor for most of our distributed experimental activities during both funding periods. The corresponding time series allowed to discriminate differences in the susceptibility for preferential flow, to assess the identification of the minimum necessary number of profiles to discriminate these differences – we found that shorter observation times go on the expense of a higher necessary amount of profiles (project G). This finding speaks clearly for long-term observatories. The sensor clusters were furthermore key to develop a slightly different perspective on spatial variability of soil moisture measurements. While spatial differences were large, corresponding ranks exhibit a remarkable temporal stability (project G). Distributed time series are in fact highly redundant and the related spatial dispersion functions/experimental variograms may be classified into temporally persistent and recurrent clusters. This strongly underpins the temporal stability and recurrence of spatial soil moisture patterns (project F). The joint data of soil water content and matric potentials were also key to calculate free energy of the water content at these sites and to inter-compare the differences among the Colpach (Schist) and the Wollefsbach (Marl.)

The cluster sites played also a key role for the earth-worm survey and related dye tracing, which turned out as a powerful means to characterise macropores (projects A and F). This revealed a clear relation between abundance of worm up to 16 different worm species and hydrologically relevant biopores and allowed derivation of time series of hydrologically relevant active macropores using species distribution models. These data allowed a successful parameterization of the Lagrangian models echORD and LAST (project F). But we have to admit that these data were not yet used for larger scale simulation studies with CATFLOW, the CAOS model or any other hydrological model.

Soil moisture time series at the cluster sites were together with locally collected stable water isotope profiles in soils most helpful to shed light on vadose zone water storage and mixing, and to characterise residence times of seepage and evaporated water in the upper 2m (project D). This data together with nearly a decade of tracer work in the Attert River basin shed new light on the mechanism of how critical zone structures affect water chemistry-discharge relationships. Pb isotopes have a great potential to serve as ideal conservative tracers – because any change in isotopic ratios is solely controlled by mixing of different sources and end member. Furthermore, we demonstrated that  $^3\text{H}$  contains valuable information that is not in common with  $\delta^2\text{H}$  when inferring transport processes to the stream (project D). Interestingly, the median residence time of water in the Weierbach headwater inferred from these two isotopic tracers was consistent with the median residence time of water in the upper 2 m of the unsaturated zone. This suggest that the unsaturated zone should get more attendance in isotope hydrology.

Using more traditional methods and free energy of the soil water content we are able to a) explain the frequently observed double peaks in the Colpach by storage threshold (project D) and b) the onset of stream flow generation by a switch between storage deficit to a storage excess in the riparian zone

(project F). Simple double mass curves yielded furthermore a similar seasonal interplay between discharge and rainfall input in the Marl and Schist geology with a strong inter-annual variability, while the sandstone acted like a low pass filter without much variation between the years (project F). This finding was well confirmed by the comprehensive analysis of distributed temperature and electric conductivity measurements in the stream network (project G).

Importantly, we developed several approaches to investigate hydrological connectivity and riparian zone exchange. While Radon did not allow the quantification of riparian zone exchange fluxes, it was well but to detect its small scale variability on the 10 m scale (project G). Terrestrial diatoms are furthermore feasible for detecting source areas of surface and subsurface runoff components (project D). Last but not least, the joint event scale analysis of foot-slope piezometer reactions and streamflow responses turned out to be very well suited for the assessment of subsurface hillslope-stream connectivity (project G).

An advanced radar based precipitation climatology was derived by fusing operational C-Band weather radar data with two MRR vertical radars, six disdrometers and seven rain gauges using a combination of geo-statistical merging in combination with information theory (project C).

Finally and yet importantly, the sap flow data collected at the cluster sites proved to be valuable to quantify tree transpiration rates in summer and to inter compare those between geologies, expositions and slopes (project G and F). Interestingly, transpiration rates in sandstone appeared larger, though soil water storage in the upper 50 cm appeared similar. This suggests that trees tap a much deeper store in sandstone. Sapflow data were also key to test the feasibility to infer root water uptake based on diurnal decline in soil water content (project F). The joint field campaign to evaluate the energy balance and ET of project C, was another highlight, which particularly revealed the feasibility of UAS with thermal infrared and optical cameras to infer patterns of latent and sensible heat, and their hysteresis. Also ASTER thermal infrared data were most valuable to characterise difference in landuse units, as proxy for different controls on the land-surface energy balance.

### **3.3 Q3: How to balance necessary model complexity and falsifiability? D3: Advancements in model concepts and new framework for model inter comparison using information theory.**

We conclude that information theory and thermodynamics provide also ideal match to quantify model complexity and minimum requirements with respect to the physical soundness of process descriptions in hydrological models. From the information perspective, we define an optimum model complexity as the most compressed spatial model structure that performs acceptable with respect to a pre-defined performance threshold. From a thermodynamic standpoint, we stick to the flux, gradient and resistance triple and recommend that any flux of mass, energy, momentum or entropy in a model should be expressed as product of a potential gradient and an inverse resistance, which characterizes the work the fluid performs when flowing through the soil, river, tree or aquifer.

Our findings revealed that dynamic dissimilarity of different landscape units such as hillslopes/ sub-catchments with respect to runoff generation or soil moisture dynamics, changes with time. Runoff generation turned out to be more complex than soil moisture dynamics, because functional similarity of the former generation within the hillslopes ensemble compiling the Colpach catchment is much more

variable in time than soil moisture dynamics. We thus conclude that the minimum necessary spatial model complexity depends on the modelling target and changes dynamically with the prevailing similarity of the target variable within functional units. Technically, we found that adaptive clustering is feasible to dynamically adjust the spatial model complexity to the prevailing dynamic similarity and this approach was shown to work independently from the question whether a model is of conceptual or physically based nature.

The thermodynamic perspective, however, reveals that the SHM is over simplified, while CATFLOW is the consistent with the concept of the flux-gradient-resistance paradigm. The spatial complexity of the latter can even be compressed to a single representative hillslope, which simulated the water balance of the 19 km<sup>2</sup> large Colpach catchment even better than the distributed ensemble of 105 hillslopes. From this, we conclude that the search for a representative topographic gradient is an effective way to simplify models without lumping distributed process concepts that satisfy the flux-gradient-resistance triple into effective process concepts and parameters. In line with this, the maximum power approach in combination with a new thermodynamic limit called the Carnot limit of a cold heat engine, allowed predictions of the land surface energy balance and evaporation partitioning in strong accordance with Eddy flux tower observations. This is remarkable, as the approach does neither make use of empirical turbulence parametrizations nor a stomata conductance model. Simulations of the WRF corroborated on the contrary that a change in the stomata conductance models in traditional soil-water-atmosphere transfer schemes have a high impact on the simulated energy balance and bower ratio. These findings suggest that canopy scale turbulent exchange of CO<sub>2</sub> and water vapor might be the limiting factors of photosynthesis during convective conditions. We thus conclude that the other key for model simplifications is a proper representation of energetic limitations.

An appropriate model complexity implies, however, also to improve in appropriate process descriptions, even if they are in line with the flux-gradient-resistance triple. In this respect we developed and tested to promising approaches to deal with preferential flow, as the Richard's equation is inappropriate here. The Lagrangian model LAST (Zehe and Jackisch, 2016; Sternagel et al., 2019) offers many assets to simulate preferential flow and reactive transport in heterogeneous soils compared to the traditional Richards and advection-dispersion equations. Similarly, the stokes flow yielded promising results for infiltration and preferential flow, also for application at larger scales (Demand et, 2019). The study of Loritz et al. (2017) revealed furthermore, that a uniform root water uptake is inappropriate to capture variability of soil water dynamics within the root zone.

Despite these achievements, we must admit that we did not live up to our goals with respect to the CAOS model development. The current model version performs clearly worse than for instance that CATFLOW model and it is less computational efficient. While we succeeded in developing an information based framework for model inter-comparison and uncertainty assessment, we did not inter compare the CATFLOW, the ROGER, CAOS and the NOAH-MP-HYDRO models within this framework as intended. This is because CAOS model simulation quality was not sufficient, and project D neither succeed in applying the ROGER nor did project C set up the NOAH-MP model for the Attert. The CATFLOW model was successfully setup for the Colpach using 105 hillslopes, and here we used information entropy to shed light on the time varying dynamic similarity of simulated runoff and storage,

as well as a new mutual information index as metric for clustering functionally similar hillslopes into functional units. Our new conceptual model SHM (Simple Hydrological Model) was the main testbed for the development of dynamical adaptive model clustering. As our framework for model evaluation turned out to be much more generally applicable than envisaged, we decided to shift from the original set of models to a more diverse range encompassing purely databased and conceptual models. This revealed that with just a few well-chosen predictors (mainly precipitation sums and gauge observations of the immediate past), purely databased models perform as good as conceptual models. While this is somewhat disenchanting for conceptual hydrological modellers, it underpins how much information about discharge is contained in both the meteorological drivers and in the memory of the system.

### **3.4 Q4: What is the connection between thermodynamics, organising principles, catchment structure and functioning? D4: Protocol for testing organizing principles.**

We finally continue our reflection on the value of thermodynamic concepts for a better understanding the catchment structure/catchment form and functioning, and constraining hydrological predictions. Project C explored a thermodynamic perspective for a better understanding of diurnal variations of surface energy balance partitioning with emphasis on solar radiation, and proposed a new thermodynamic limit called *the Carnot limit of a cold heat engine* (Kleidon and Renner, 2018). The latter allowed improved predictions of the land surface energy balance and evaporation partitioning by assuming maximum power in the sensible heat flux at different land covers (Conte et al., 2019). This approach was further refined by including diurnal changes in the storage in soil and the boundary layer (Kleidon and Renner, 2018). Overall, these findings corroborate that thermodynamics provides an additional, relevant constraint to infer energy balance partitioning and evaporation, which essentially requires a systems perspective. This is because interactions with surface temperature set the associated thermodynamic limit, this approach reproduces large-scale patterns of energy balance partitioning and the hydrologic budget as well as the diurnal cycle, and it adequately captures sensitivities to global change (Kleidon and Renner, 2017).

Complementary to this, projects F (and S) developed a new thermodynamic perspective on soil water dynamic dynamics and runoff generation, because water flows and changes in water stocks relate to conversion and mostly dissipation of gravitational and capillary potential energies and associated production of entropy. A related new system characteristics was shown to be well suited to distinguish a) state ranges corresponding either to a storage excess or a storage deficit and b) to differences in how gravity and capillary joint control on soil water dynamics between the Colpach and the Wollefsbach (Zehe et al., 2019). The free energy state of soil water in the riparian zone of both study catchments has furthermore been proven to be helpful to explain the threshold like onset in streamflow generation, which coincided with the transition from a storage deficit to a storage excess. Additionally, Loritz et al. (2019) proposed a novel thermodynamic index  $r$ -dune, in fact the ration of the driving potential energy difference to the dissipative losses along the flow path, which explains the observed differences in runoff generation between the three geologies of the Attert clearly better than HAND or the well-known topographic index.

The energy perspective is also helpful to explain the ubiquitous preferential flow phenomena. Preferential flow in connected networks implies a more energy efficient throughput of water and matter through the system. This because they reduce flow-weighted dissipative losses due to an increased

hydraulic radius in the rill or river network compared to sheet overland flow (Howard, 1990; Kleidon et al., 2013) or in subsurface connected preferential pathways compared to matrix flow (Hergarten et al., 2014; Zehe et al., 2010). At the end of the day, water prefers the path of minimum resistance, due to the highly dissipative nature of porous media and shallow overland flow. The tradeoff between energy dissipation via overland flow or via preferential infiltration, implies also the existence of an optimum soil infiltration capacity maximizing total energy dissipation. While we do not claim this to be a universal constrain on rainfall splitting into overland flow and infiltration, it provides at least a testable hypothesis for uncalibrated predictions (Zehe et al., 2013). Despite these findings we did so far not test whether this optimality approach can be used to constrain simulations of the water balance in the Wollefsbach (Loritz et al., 2017). This is because we found that formation of shrinkage cracks in summer dominates the runoff generation, and this process is so far not included in the CATFLOW model.

The concept of energy conservation was also most helpful to model simplification and upscaling. We demonstrated that the water balance of the 19 km<sup>2</sup> large catchment can be simulated using a single 2d effective hillslope model (Loritz et al., 2017). The key was to derive effective topography such that the average distribution of potential energy along the average flow path length to the stream was conserved. Similarly, we derived a macroscale effective soil water retention curve such that it conserved the relation between the average soil water content and matric potential energy using a set scale retention experiments with soil cores (Jackisch 2015; Zehe et al., 2019).

We thus conclude that the combination of a thermodynamic and information centred perspective on hydrological systems with process research offers new and promising perspectives for explaining and comparing hydrological dynamics and similarity, new avenues in modelling (including model simplification) and uncertainty assessment, which cannot be inferred from traditional water balance thinking alone.

#### 4 Publication of data from final reports

Sprenger, Matthias; Seeger, Stefan; Blume, Theresa; Weiler, Markus (2021): Stable isotope (<sup>2</sup>H and <sup>18</sup>O) depth profiles of pore waters and inferred soil physical parameters in the Attert catchment, Luxembourg. V. 1. GFZ Data Services. <https://doi.org/10.5880/fidgeo.2021.030>.

Keller, Sina; Riese, Felix M.; Allroggen, Niklas; Jackisch, Conrad (2020): HydReSGeo: Field experiment dataset of surface-sub-surface infiltration dynamics acquired by hydrological, remote sensing, and geophysical measurement techniques. GFZ Data Services. <https://doi.org/10.5880/fidgeo.2020.015>

Jackisch, Conrad; Hassler, Sibylle K. (2019): Rhizosphere soil moisture dynamics and sap flow – determining root water uptake in a case study in the Attert catchment in Luxembourg. GFZ Data Services. <https://doi.org/10.5880/fidgeo.2019.030>

van Schaik, Nicolette Loes M.B.; Zangerlé, Anne; Hohenbrink, Tobias L.; Reck, Arne; Schneider, Anne-Kathrin; Schröder, Boris (2019): Spatiotemporal variability in infiltration through biopores: earthworms, macropores and infiltration patterns. V. 1.1. GFZ Data Services. <https://doi.org/10.5880/fidgeo.2019.022>

Neuper, Malte; Ehret, Uwe (2018): CAOS rain rate and reflectivity data set of 6 disdrometres and 2 micro rain radars at 3 different heights at 6 stations in the Attert catchment, Luxembourg from Oct 2012 - Sept 2016. GFZ Data Services. <https://doi.org/10.5880/FIDGEO.2018.070>

Renner, Maik; Wizemann, Hans-Dieter; Brenner, Claire; Mallick, Kaniska; Trebs, Ivonne; Wulfmeyer, Volker; Schulz, Karsten; Kleidon, Axel (2018): Surface energy balance at a grassland site in Luxembourg modelled by three structurally different evapotranspiration schemes. GFZ Data Services. <https://doi.org/10.5880/fidgeo.2018.019>

Wizemann, Hans-Dieter; Trebs, Ivonne; Wulfmeyer, Volker (2018): Surface energy balance observations at a grassland site in Luxembourg. GFZ Data Services. <https://doi.org/10.5880/fidgeo.2018.024>

#### 5 References not included in section 1.4

Beven, K., and Germann, P.: Macropores And Water-Flow In Soils, *Water Resources Research*, 18, 1311-1325, 1982.

Beven, K. and Germann, P., Macropores and water flow in soils revisited, *Water Resour. Res.*, 49(6), 3071–3092, doi:10.1002/wrcr.20156, 2013.

Cenci L., Pulvirenti L., Boni G., Chini M., Matgen P., et al.. An evaluation of the potential of Sentinel 1 for improving flash flood predictions via soil moisture–data assimilation. *Adv. Geosci.* 44: 89–100, 2017.

Davies, J., Beven, K., Rodhe, a., Nyberg, L. and Bishop, K., Integrated modeling of flow and residence times at the catchment scale with multiple interacting pathways, *Water Resour. Res.*, 49(8), 4738–4750, doi:10.1002/wrcr.20377, 2013.

- Dietrich, W. E., and Perron, J. T.: The search for a topographic signature of life, *Nature*, 439, 411 – 418, 2006.
- Dudhia, J.: Numerical study of convection observed during winter monsoon experiment using a mesoscale twodimensional model. – *J. Atmos. Sci.* 46, 3077–3107, 1989.
- Ewen, J.: “SAMP” model for water and solute movement in unsaturated porous media involving thermodynamic subsystems and moving packets. 1. Theory, *J. Hydrol.*, 182(1–4), 175–194, doi:10.1016/0022-1694(95)02925-7, 1996.
- Ewen, J.: ‘SAMP’ model for water and solute movement in unsaturated porous media involving thermodynamic subsystems and moving packets: 2. Design and application, *J. Hydrol.*, 182(1–4), 195–207, doi:10.1016/0022-1694(95)02926-5, 1996.
- Farquhar, G. D., S. von Caemmerer, and J. A. Berry, A biochemical model of photosynthetic CO<sub>2</sub> assimilation in leaves of C<sub>3</sub> species, *Planta*, 149(1), 78–90. doi:10.1007/BF00386231, 1980.
- Hergarten, S., G. Winkler, and S. Birk: Transferring the concept of minimum energy dissipation from river networks to subsurface flow patterns, *Hydrology And Earth System Sciences*, 18(10), 4277-4288, 2014.
- Howard, A. D.: Theoretical model of optimal drainage networks, *Water Resour. Res.*, 26, 2107-2117, 1990.
- Hincapie, I. A., and Germann, P. F.: Abstraction from Infiltrating Water Content Waves during Weak Viscous Flows, *Vadose Zone Journal*, 8, 996-1003, 10.2136/vzj2009.0012, 2009.
- Jackisch, C., Linking structure and functioning of hydrological systems., PhD. Thesis, KIT - Karlsruhe Institute of Technology., doi: 10.5445/IR/1000051494, 2015.
- Pfister, L., Humbert, J., Iffly, J.F., Hoffmann, L.: Use of regionalized stormflow coefficients in view of hydro-climatological hazard mapping. *HSJ* 47: 479-491, 2002.
- Pfister, L., McDonnell, J. J., Hissler, C., and Hoffmann, L.: Ground-based thermal imagery as a simple, practical tool for mapping saturated area connectivity and dynamics, *Hydrological Processes*, 24 (21), 3123–3132, doi:10.1002/hyp.7840, 2010.
- Simunek, J., Jarvis, N. J., van Genuchten, M. T. and Gardenas, A.: Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone, *J. Hydrol.*, 272, 14–35, doi:10.1016/S0022-1694(02)00252-4, 2003.
- Shipitalo MJ, Butt KR.: Occupancy and geometrical properties of *Lumbricus terrestris* L. burrows affecting infiltration. *Pedobiologia* 43: 782-794, 1999.
- Sternagel, A., Loritz, R., Klaus, J., Berkowitz, B., and Zehe, E.: Simulation of reactive solute transport in the critical zone: a Lagrangian model for transient flow and preferential transport, *Hydrology and Earth System Sciences*, 25, 1483-1508, 10.5194/hess-25-1483-2021, 2021.
- Strunck, L.: Unsicherheitsabschätzung in der Skalierung von Safflussmessungen anhand eines Datensatzes aus dem Atert-Einzugsgebiet in Luxemburg. BSc.-Thesis at the Institute for Water and River Basin Management, KIT, 51 pp, 2020.

- Sun, J., N.A. Crook: Dynamical and microphysical retrieval from doppler radar observations using a cloud model and its adjoint. Part I: Model development and simulated data experiments. *J. Atmos. Sci.* 54, 1624–1661, 1997.
- Thiem, Christina Elisabeth, Matthias Mauder, Christian Chwala, Matthias Bernhardt, Harald Kunstmann, and Karsten Schulz. “Intra-Sensor Variability Study of Two BLS 900 Scintillometers.” In *EGU General Assembly Conference Abstracts 19*, 1677, 2017.
- Troch, P. A., G. Carrillo, M. Sivapalan, T. Wagener, and K. Sawicz Climate-vegetation-soil interactions and long-term hydrologic partitioning: signatures of catchment co-evolution, *Hydrology And Earth System Sciences*, 17(6), 2209-2217, 2013.
- Xiao, Q., J. Sun: Multiple-radar data assimilation and short-range quantitative precipitation forecasting of a squall line observed during IHOP\_2002. – *Mon. Wea. Rev.* 135, 3381–3403, 2007.
- Weiler M.: Mechanisms controlling macropore flow during infiltration. PhD-Thesis. Technischen Wissenschaften ETH Zürich Nr. 14237. DOI:10.3929/ethz-a-004180115, 2001.
- Zehe, E., and Fluhler, H.: Preferential transport of isotoproturon at a plot scale and a field scale tile-drained site, *Journal of Hydrology*, 247, 100-115, 2001.
- Zehe, E., and Blöschl, G.: Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions, *Water Resources Research*, 40, W1020210.1029/2003wr002869, 2004.