Heterogeneous Land-Surface Effects on TKE and Cloud Formation: Statistical Insights from LES Cases

Jason S. Simon[1,2], Andrew D. Bragg[1], and Nathaniel W. Chaney[1]

[1]Duke University, Durham, NC, USA
[2]Present Address: Saint Augustine's University, Raleigh, NC, USA

Key Points:

- Analysis of 92 LES cases shows strong statistical correlations between land-surface heterogeneity and mesoscale atmospheric development
- Correlation between surface heterogeneity and circulation/cloud production is dominated by the largest wavelengths in the land-surface field

Corresponding author: Jason Simon, `jasonsimon@berkeley.edu`

Abstract
To aid development of sub-grid scale (SGS) parameterizations for Earth system models which consider heterogeneity in land-surface fields and land-atmosphere coupling, results from large-eddy simulations of 92 shallow convection cases over the Southern Great Plains are presented and analyzed. Each case is simulated with heterogeneous surface fields obtained from an offline field-scale land-surface model, and with spatially homogeneous surface fields with the same domain-wide mean value. By comparing corresponding heterogeneous and homogeneous cases, it is found that turbulent kinetic energy and liquid water path has a high correlation with the spatial variance of the surface heat flux fields. By further comparing the source of this correlation over the range of wavelengths in the surface fields, it is found that the majority of the heterogeneous land-atmosphere coupling is contained in wavelengths of order 10 km and larger, suggesting an encouraging degree of feasibility of including land-surface heterogeneity in global-scale SGS parameterizations.

Plain Language Summary

To help efforts to alleviate some of the issues associated with the relatively low-resolution grids used by modern global weather and climate models, we first created a dataset of 92 high-resolution simulations over the Southern Great Plains region of Oklahoma. All of the cases in the dataset are based on days which where observed to produce shallow clouds, which can have a significant impact on the incoming solar radiation. The high-resolution simulations were designed to cover a region large enough to contain relevant cloud production which is also too small to be represented on a modern global model. The dataset of high-resolution simulations is analyzed to compare the strength of the patterns in the land surface to the associated increase in cloud production. It is hoped that this and similar future studies will provide insights which increase the fidelity of cloud production models which intend to capture effects which are smaller than the grid used for global models.

1 Introduction

Modern coupled Earth system models (ESMs) are run at horizontal resolutions that are $\mathcal{O}(10-100\text{ km})$, which is decided by the balance between computational resources and the demands of the atmospheric component of the coupled model, while the land-surface model (LSM) component could conceivably have an effective horizontal spatial resolution around $\mathcal{O}(10-100\text{ m})$ (e.g., Chaney et al., 2018). This loss of land-surface information is made more significant by the fact that it spans the relevant length scales for many important coupled processes, namely those related to boundary-layer growth and cloud production (Bertoldi et al., 2013; Kang & Bryan, 2011; Ntelekos et al., 2008; Weaver, 2004).

The parameterization associated with sub-grid scale (SGS) cumulus production is very important in contemporary ESMs, by virtue of the importance of cloud production to the Earth system in general. Many modeling and observational studies find that secondary circulations induced by thermal surface heterogeneity can act as sources of convection and significantly alter local cloud production rates and distribution (e.g., Albertson et al., 2001; Dixon et al., 2013; Kang, 2020; Marsham et al., 2008; Mendes & Prevedello, 2020; Taylor et al., 2011; Phillips & Klein, 2014).

While the aforementioned land-surface patterns are SGS on grids used for most modern global models, there is a large amount of information available regarding the characteristics of the land-surface which could potentially be utilized by SGS parameterizations. Towards this effort, we present a large-eddy simulation (LES) study of 92 shallow convection cases over the Southern Great Plains (SGP) site, based on cases developed by the LES ARM Symbiotic Simulation and Observation Workflow (LASSO) campaign (W. Gustafson et al., 2019; W. I. Gustafson et al., 2020). The cases are run using high-resolution spatially-heterogeneous land-surface fields and also using spatially-homogeneous land-surface fields, which match the heterogeneous cases' domain-wide mean values through time but contain none of the spatial structure.

We find that there is a strong correlation between basic metrics of heterogeneity in the surface heat flux fields and the resultant additional production of liquid water path (LWP) and circulating kinetic energy. We also find that, for the cases considered here, the majority of the relevant information about the heterogeneity of the land-surface is contained in the few Fourier modes of the fields with the largest wavelengths, which is encouraging from the perspective of computational resources potentially required to consider SGS land-surface features.

2 Model description

Large-eddy simulations are conducted using version 3.8.1 of the WRF model (Skamarock et al., 2008) with modifications as described by J. S. Simon et al. (2021). Cases here use a horizontal resolution of 250 m and a domain of 130×130 km^2 laterally. The land-surface fields in the outer 15 km of the domain are tapered to linearly approach their domain-wide mean on each boundary to eliminate discontinuities in the land-surface that may otherwise be introduced by the periodic boundary conditions. Each domain is also rotated to closer align the bulk liquid-water flux normally to the boundaries, based on results from an initial simulation using the unrotated land-surface, to limit artificial spreading of liquid water caused by the fluxes through the boundaries not aligning with the periodicity of the domain. The model configuration is otherwise the same as in J. S. Simon et al. (2021).

Each case is run with heterogeneous and homogeneous land-surface fields (sensible heat flux, latent heat flux, skin temperature, albedo, and momentum drag coefficient), where homogeneous cases specify a uniform (in space) surface of each field to match the time-evolving domain-wide mean of the corresponding heterogeneous case. There is no feedback from the atmosphere to the land surface in the LES; the HydroBlocks LSM is run offline and the output surface fields are specified as the bottom boundary in the WRF model. Further details of the HydroBlocks LSM and its coupling to the WRF model can be found in the Supporting Information.

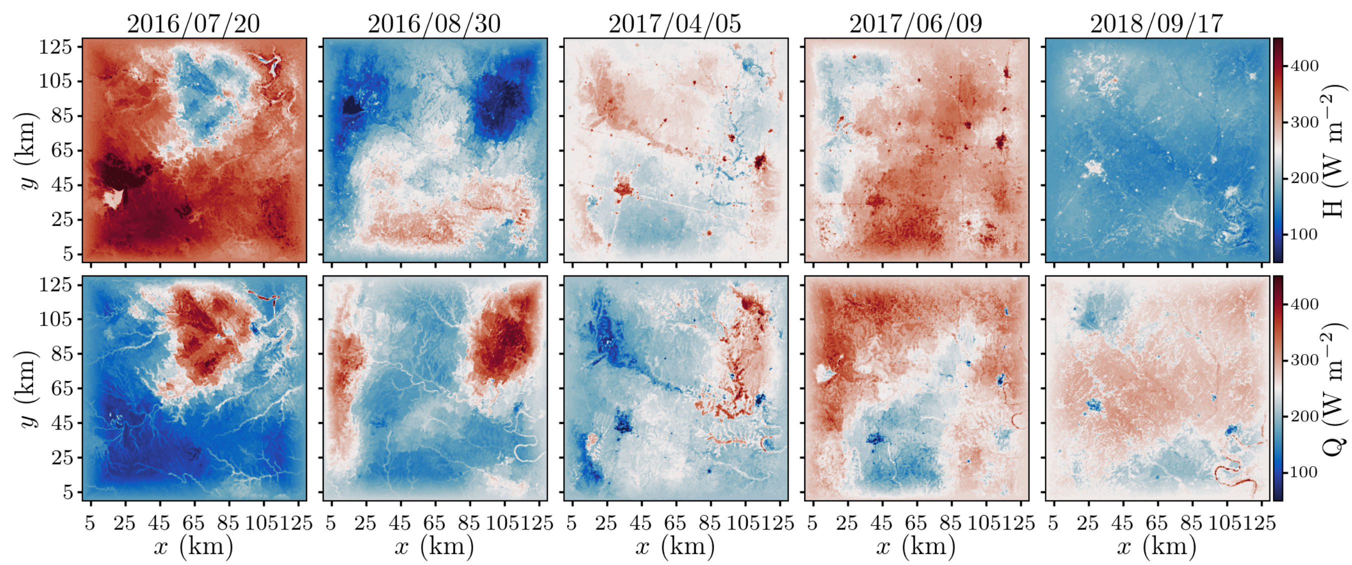

Figure 1. Example maps of sensible (H) and latent (Q) heat flux fields.

3 Results

3.1 Evaluation Metrics

The domain-wide measure of vertically-integrated, mass-coupled turbulent kinetic energy (TKE) is compared between cases, serving as a metric for general activity in ABL development. For brevity, "TKE" will refer to the vertically-integrated, mass-coupled form unless otherwise stated. On the discretized WRF grid, the TKE is found as

$$\mathrm{TKE} = \sum_z \rho_a \left[\frac{1}{2}\left(u'^2 + v'^2 + w'^2\right)\right] \Delta_z, \tag{1}$$

where ρ_a is air density, (u, v, w) are the velocity components in the (x, y, z) directions, Δ_z is the grid spacing in the vertical direction, and a primed variable indicates deviation from the mean value in the (x, y) plane. For illustration, a time series of TKE for heterogeneous and homogeneous simulations of an example case is shown in Fig. 2a. Cases are also compared by their domain-wide LWP signal, which serves as a proxy for overall cloud production. On the discretized WRF grid, our measure of LWP is found as

$$\mathrm{LWP} = \sum_z \rho_a q_l \Delta_z, \tag{2}$$

where q_l is liquid water mixing ratio.

Part of the LASSO modification to the WRF code is the addition of output solution fields as average values over a given interval of time, in addition to the standard instantaneous output fields. Here, time-averaged fields are found over 10 min intervals from samples taken internally every 30 s. Notationally, we will use $\mu(\phi)$ and $\sigma(\phi)$ to indicate the spatial mean and standard deviation, respectively, of a field $\phi = \phi(x, y)$ at a point in time. For temporal averages, we will use the notation $\mathrm{mean}[\vartheta]$, found as

$$\mathrm{mean}[\vartheta] = \frac{\sum_t g_s \vartheta}{\sum_t g_s}, \tag{3}$$

where $\vartheta = \vartheta(t)$ is a domain-wide scalar, with $g_s = g_s(t)$ defined as

$$g_s(t) = \begin{cases} 1 & : \quad s\,(t) > 0.05 \, \max(s), \\ 0 & : \quad s\,(t) \leq 0.05 \, \max(s), \end{cases} \tag{4}$$

where $s(t)$ is the surface downward clear-sky shortwave radiation at time t, and $\max(s)$ is the maximum value of s over the given simulation. The averaging procedure in (3) is used for both the atmospheric fields (Fig. 2c) and the land-surface statistics (Fig. 2e).

The heterogeneous vs. homogeneous statistics for TKE and LWP are compared using the metric $\gamma(\vartheta)$, defined as

$$\gamma(\vartheta) = \mathrm{mean}\left[\log \gamma_t(\vartheta)\right], \tag{5}$$

where

$$\gamma_t(\vartheta) = \frac{g_s \vartheta_{\text{heterogeneous}} + 1}{g_s \vartheta_{\text{homogeneous}} + 1}. \tag{6}$$

Equations (6) and (5) are demonstrated visually in Fig. 2c and d, respectively. The form of (6) is motivated as a ratio of ϑ between heterogeneous and homogeneous cases, which is weighted by g_s to isolate daytime values. The addition of 1 to both terms is included to limit the influence of very small values which are effectively negligible, as well as to avoid the edge cases of $\gamma_t = 0$ or $\gamma_t = \infty$ when $\vartheta_{\text{heterogeneous}} = 0$ or $\vartheta_{\text{homogeneous}} = 0$, respectively. When $\vartheta \approx 0$ for both the heterogeneous and homogeneous cases, $\gamma_t \approx 1$, indicating the two cases have approximately equal measures of ϑ, as intended. In (6), TKE is given in units of kg s^{-2} and LWP in units of g m^{-2}.

In addition to γ(TKE) and γ(LWP), we also compare corresponding heterogeneous and homogeneous cases by only the circulating portion of kinetic energy. This is found from the turbulence spectra of 10-minute averaged u and v fields, where only the energy from modes that are in the lowest 5 km of the atmosphere and longer than 10 km laterally are included. The ratio of circulating energy between the heterogeneous and homogeneous cases, which we denote χ, is found as similarly to the TKE and LWP fields, as

$$\chi = \text{mean}\left[\log \sqrt{\gamma_t(E_u)\gamma_t(E_v)}\right], \tag{7}$$

where

$$E_\varphi = E_\varphi\left(t\right) = \sum_{z<5\text{ km}} \left[\sum_{\ell>10\text{ km}} |\widehat{f}(\varphi')|^2\right], \tag{8}$$

and $|\widehat{f}(\varphi')|$ is absolute value of the normalized two-dimensional discrete Fourier transform of φ', ℓ is the component of the Fourier mode's wavelength, λ, in the direction aligned with φ (e.g., $\ell = \lambda_x$ for $\varphi = u$).

Two length-scale metrics, L_Δ and L_2, are presented for the land-surface fields, based on their Fourier spectra (the relaxation to the mean value on the outer 15 km of the land-surface fields render their boundaries as effectively periodic). The L_Δ length scale gives the approximate scale of the largest coherent structures in the field, and is found as

$$L_\Delta(\phi) = \frac{\sum_\lambda \lambda \Delta_\lambda \sqrt{|\widehat{f}(\phi)|}}{\sum_\lambda \Delta_\lambda \sqrt{|\widehat{f}(\phi)|}}, \tag{9}$$

where $\phi = \phi(x, y)$ is a heterogeneous surface field, and Δ_λ is the difference between λ and the next (smaller) wavelength in the discrete spectrum. The L_2 length scale gives the approximate scale of the smallest coherent structures in the field, and is found as

$$\left[L_2(\phi)\right]^2 = \frac{\sum_\lambda \lambda^2 \sqrt{|\widehat{f}(\phi)|}}{\sum_\lambda \sqrt{|\widehat{f}(\phi)|}}. \tag{10}$$

Correlations between atmosphere and land-surface fields are evaluated by the Pearson (ρ_p) and Spearman (ρ_s) correlation coefficients, as implemented by Virtanen et al. (2020) (e.g., Fig. 2f).

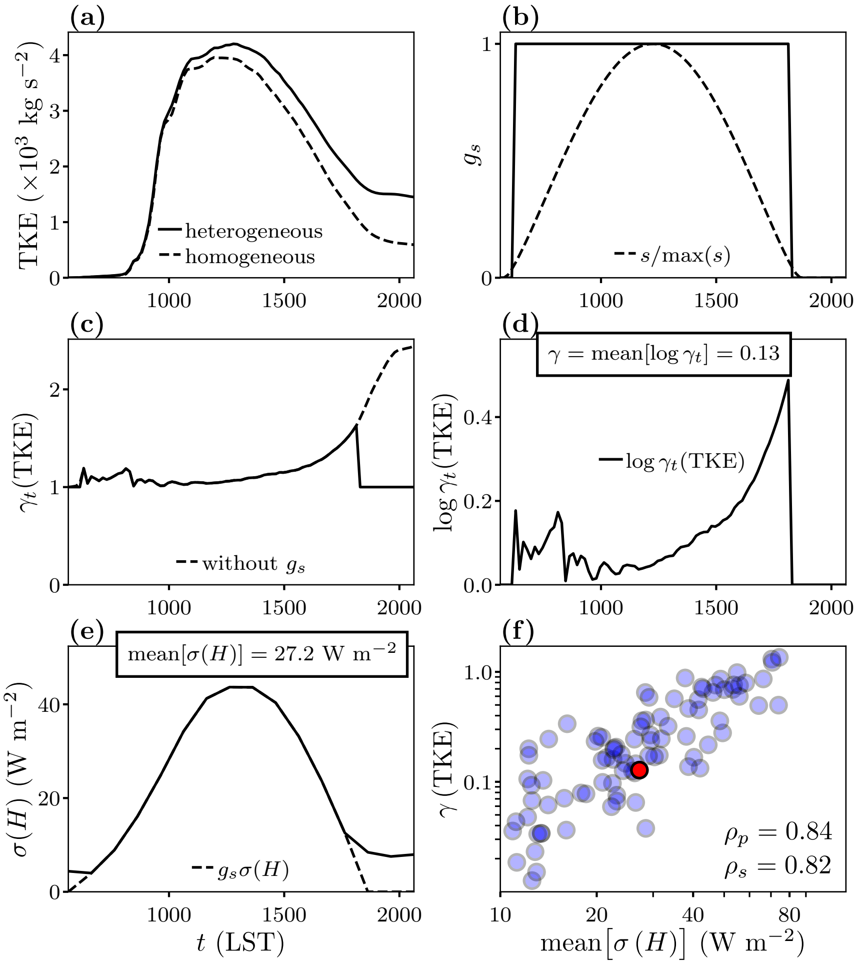

Figure 2. Demonstration of the comparison process for heterogeneous and homogeneous cases: (a) the domain-wide time series of TKE for the two simulations of 2017/08/30; (b) the time filter, g_s, used by the mean[ϑ] function as described by (4); (c) the calculation of γ_t, as described by (6); (d) the calculation of γ as described by (5); (e) the application of the time filter, g_s, and the mean[ϑ] function to the $\sigma(H)$ time series; (f) an example scatter plot of mean[$\sigma(H)$] vs. γ(TKE) for all 92 days with the datapoint for 2017/08/30 shown in red.

3.2 Land-Atmosphere Correlation

The emergent secondary circulations driven by land-surface heterogeneity are analyzed by the correlation between χ and each of $\mu(H)$, $\sigma(H)$, $L_\Delta(H)$, and $L_2(H)$ (Figs. 3a–d, respectively). There is a strong positive relationship between χ and all of $\sigma(H)$, $L_\Delta(H)$, and $L_2(H)$, but only a trivial correlation with $\mu(H)$. The same presentation is repeated for $\gamma(\text{LWP})$ in Fig. 3e–h. The $\gamma(\text{LWP})$ data is very similar to that of the χ metric, but with $\sim 20\%$ smaller magnitudes. Of the 92 cases, 4 have more liquid water production in the homogeneous simulation, indicated by a negative value of $\gamma(\text{LWP})$; these datapoints are not shown in Fig. 3 but are included in the calculation of the correlation coefficients.

Visually, the data for $\gamma(\text{LWP})$ compared to $\sigma(H)$, $L_\Delta(H)$, and $L_2(H)$ show a very similar pattern as χ but with a broader spread, suggesting from that LWP production is statistically driven similarly to circulation production but with additional considerations which are not captured by the land-surface heterogeneity, which is certainly in agreement with the physical perspective of ABL development. The same analysis considering $\gamma(\text{TKE})$, or using statistics from the latent heat flux or skin temperature fields gives very similar results, which is presented in the Supporting Information.

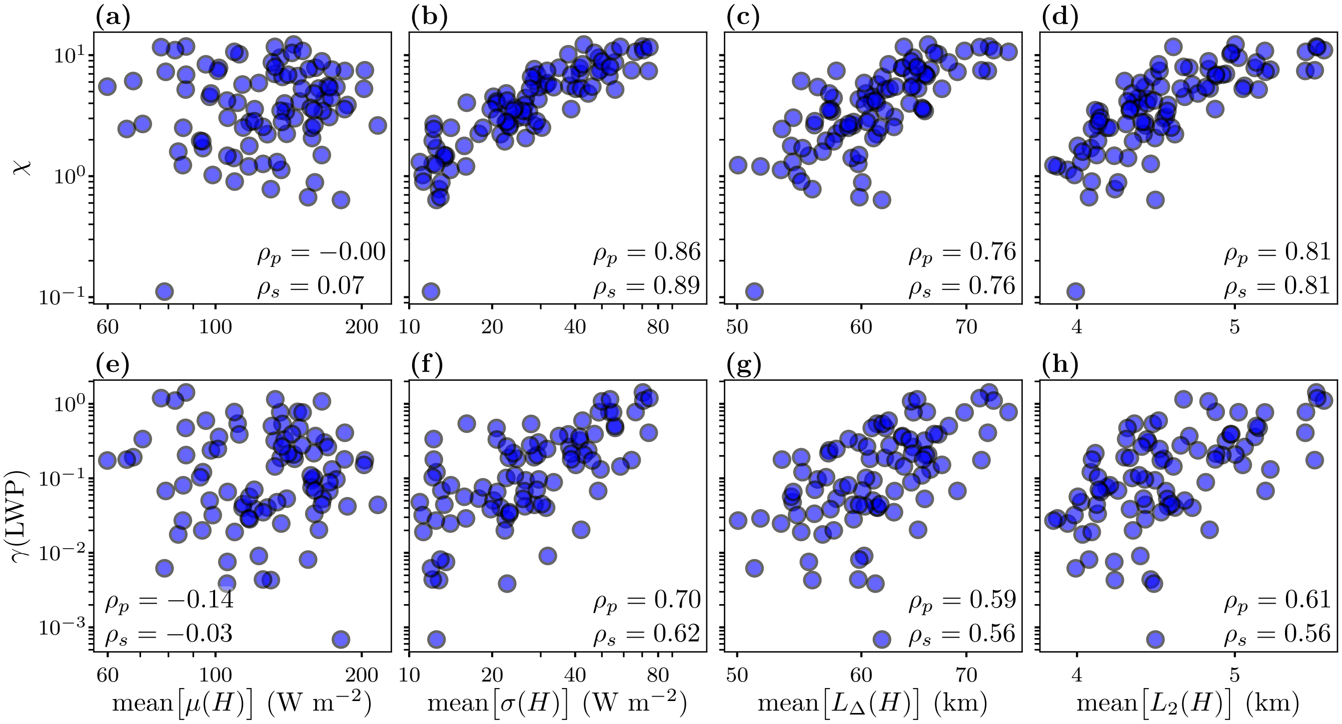

Figure 3. The χ (a – d) and γ(LWP) (e – h) metrics as functions of statistics of the surface sensible heat flux field, H. Four negative-valued data points for γ(LWP) with magnitudes $\mathcal{O}(10^{-4})$ are not shown, but are included in the calculation of ρ_p and ρ_s.

3.3 Dominant Length Scales

The results presented in Sec. 3.2 bolster the motivation to include information about the underlying land-surface heterogeneity in global SGS boundary-layer parameterizations. The best methodology to make such considerations in either existing or new parameterization models is not immediately obvious, and potential solutions must add value on a level that is commensurate with their computational and implementation costs. To evaluate the relevance of the different ranges of length scales present in the land surface on the present dataset, the land surface fields are filtered over a range of length scales and compared. The filter, F, is applied in Fourier space as

$$F = 1 - \exp\left[-2\pi^2\left(\frac{\Delta_{\text{filter}}}{\lambda}\right)^2\right], \tag{11}$$

where Δ_{filter} is the nominal filter length. To avoid ambiguity in the discussion, the operation of F is referred to as passing $\lambda < \Delta_{\text{filter}}$. An example of a mid-day sensible heat flux field from the dataset for different filter lengths is shown in the Supporting Information.

For each filter length, the heterogeneous land-atmosphere coupling is reevaluated following the same procedure as in Sec. 3.2. The average value over the dataset of mean$[\sigma(H)]$ as a function of filter length is shown in Fig. 4a. Correlation coefficients for mean$[\sigma(H)]$ of the filtered dataset with γ(TKE), γ(LWP), and χ are shown as a function of filter length in Fig. 4b, c, and d, respectively.

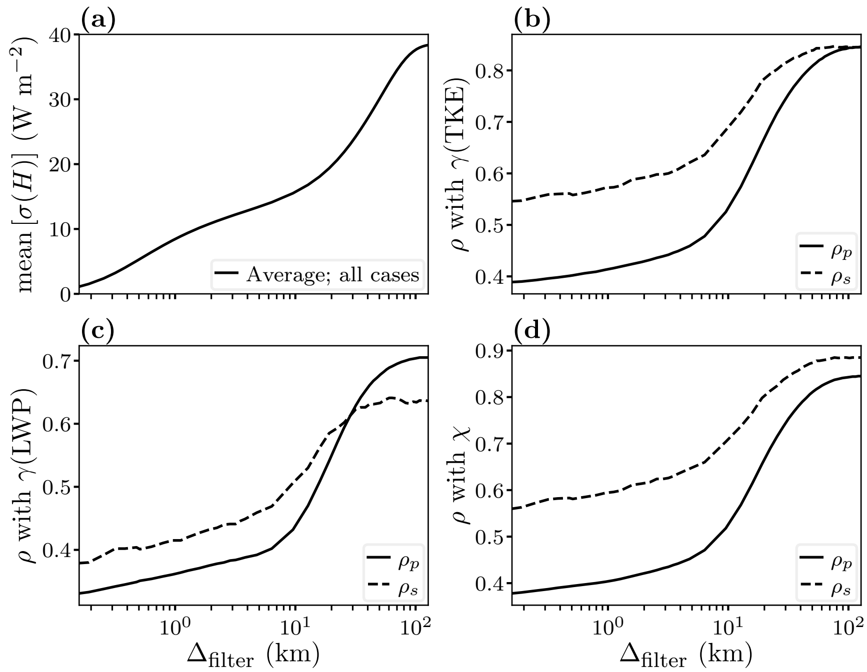

Figure 4. Average value of mean[$\sigma(H)$] over the 92 days after filtering (a) and correlation coefficients of mean[$\sigma(H)$] after filtering with γ(TKE) (b), γ(LWP) (c), and χ (d).

Figure 4a demonstrates that the majority of the standard deviation in the sensible heat flux field over the dataset is contained in length scales 10 km and larger. Figures 4b – d show the same concentration at length scales larger than 10 km for the Pearson and Spearman correlation coefficients between $\sigma(H)$ and γ(TKE), γ(LWP), χ. Because the wavelengths of Fourier modes grow geometrically, the results seen in Fig. 4 suggest that the bulk of the correlation between $\sigma(H)$ and the atmospheric metrics is contained in the longest few modes.

4 Discussion and Conclusions

We have presented a statistical analysis of the TKE and cloud production caused by land-surface heterogeneity for 92 LES cases representing different summer days from 2015 – 2019 over the SGP site by comparing simulations using heterogeneous and homogeneous land-surface fields. In Sec. 3.2 it is found that, despite all 92 days having unique initial profiles and large-scale tendencies, there is a strong correlation between the production of circulating TKE (measured as the metric χ) over a diurnal cycle and land-surface heterogeneity. The correlation between cloud production, as measured by LWP, is $\sim 20\%$ smaller but is also significant. It is also seen in Sec. 3.3 that a large portion of the correlation between the atmosphere and heterogeneous land-surfaces is concentrated in a relatively small number of the largest (by wavelength) modes in the land-surface fields.

The results in Sec. 3.2 demonstrate a strong, but incomplete, correlation between heterogeneous land surface fluxes and secondary circulations. The land-surface heterogeneity is more strongly related to χ than LWP, which was expected: while TKE production does depend on the temperature and stability of the initial atmospheric profile, liquid water production is additionally constrained by condensation conditions. Still, the correlation coefficient values seen between γ(LWP) and $\sigma(H)$, even without considerations for the state of the atmosphere, are quite strong with $\rho_p = 0.70$ (Fig. 3f). The concentration of relevant land-surface heterogeneity in structures with length scales of $\mathcal{O}(10 \text{ km})$ and larger seen in Sec. 3.3 is easily understood in the context of heterogeneous land-atmosphere coupling being largely driven by emergent mesoscale circulations. That there is such a sharp increase in correlation contained in the longest few modes of the land surface does have the encouraging implication that the level of detail necessary for the successful development of global-scale SGS parameterizations of heterogeneous land-atmosphere coupling may not be overwhelming.

While a large amount of additional work is necessary before the realization of an effective parameterization, the results seen here are encouraging. The most immediate future work is a detailed analysis of the relationship between initial and large-scale atmospheric conditions and land-surface heterogeneity on the atmospheric response. The necessary increase in cases to realize such an experiment would also enable the use of more sophisticated methods for analysis, perhaps eventually including machine learning, which itself has the potential to provide a huge value to parameterization development efforts.

Acknowledgments

Funded by NOAA grant NA19OAR4310241.

Data Availability

Simulations here use a modification of WRF version 3.8.1 developed and maintained by the LASSO team. The base WRF code, initial sounding files, and large-scale forcing files are available from W. Gustafson et al. (2019). Additional modifications to the WRF code to specify heterogeneous surfaces, data files for surface fields for each simulation, and model control files for each simulation are available at J. Simon et al. (2023a). Relevant model output and scripts used for analysis and plotting are available J. Simon et al. (2023b).

References

Albertson, J. D., Kustas, W. P., & Scanlon, T. M. (2001). Large-eddy simulation over heterogeneous terrain with remotely sensed land surface conditions. *Water Resources Research*, *37*(7), 1939–1953.

Bertoldi, G., Kustas, W. P., & Albertson, J. D. (2013). Evaluating source area contributions from aircraft flux measurements over heterogeneous land using large-eddy simulation. *Boundary-layer meteorology*, *147*(2), 261–279.

Chaney, N. W., Van Huijgevoort, M. H., Shevliakova, E., Malyshev, S., Milly, P. C., Gauthier, P. P., & Sulman, B. N. (2018). Harnessing big data to rethink land heterogeneity in Earth system models. *Hydrology and Earth System Sciences*, *22*(6), 3311–3330.

Dixon, N., Parker, D., Taylor, C., Garcia-Carreras, L., Harris, P., Marsham, J., ... Woolley, A. (2013). The effect of background wind on mesoscale circulations above variable soil moisture in the sahel. *Quarterly Journal of the Royal Meteorological Society*, *139*(673), 1009–1024.

Gustafson, W., Vogelmann, A., Cheng, X., Dumas, K., Endo, S., Johnson, K., ... Xiao, H. (2019). Description of the LASSO data bundles product. DOE Atmospheric Radiation Measurement (ARM) user facility. *DOE/SC-ARM-TR-216*.

Gustafson, W. I., Vogelmann, A. M., Li, Z., Cheng, X., Dumas, K. K., Endo, S., ... Xiao, H. (2020). The Large-eddy simulation (LES) Atmospheric Radiation Measurement (ARM) Symbiotic Simulation and Observation (LASSO) activity for continental shallow convection. *Bulletin of the American Meteorological Society*, *101*(4), E462–E479.

Kang, S.-L. (2020). Effects of mesoscale surface heterogeneity on the afternoon and early evening transition of the atmospheric boundary layer. *Boundary-Layer Meteorology*, *174*(3), 371–391.

Kang, S.-L., & Bryan, G. H. (2011). A large-eddy simulation study of moist convection initiation over heterogeneous surface fluxes. *Monthly weather review*, *139*(9), 2901–2917.

Marsham, J. H., Parker, D. J., Grams, C. M., Johnson, B. T., Grey, W. M., & Ross, A. N. (2008). Observations of mesoscale and boundary-layer scale circulations affecting dust transport and uplift over the sahara. *Atmospheric Chemistry and Physics*, *8*(23), 6979–6993.

Mendes, C. B., & Prevedello, J. A. (2020). Does habitat fragmentation affect landscape-level temperatures? A global analysis. *Landscape Ecology*, *35*(8), 1743–1756.

Ntelekos, A. A., Smith, J. A., Baeck, M. L., Krajewski, W. F., Miller, A. J., & Goska, R. (2008). Extreme hydrometeorological events and the urban environment: Dissecting the 7 july 2004 thunderstorm over the Baltimore MD Metropolitan Region. *Water Resources Research*, *44*(8).

Phillips, T. J., & Klein, S. A. (2014). Land-atmosphere coupling manifested in warm-season observations on the US southern great plains. *Journal of Geophysical Research: Atmospheres*, *119*(2), 509–528.

Simon, J., Chaney, N., & Bragg, A. (2023a, August). *Data for: Heterogeneous Land-Surface Effects on TKE and Cloud Formation: Statistical Insights from LES Cases (model and input).* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.8240267` doi: 10.5281/zenodo.8240267

Simon, J., Chaney, N., & Bragg, A. (2023b, August). *Data for: Heterogeneous Land-Surface Effects on TKE and Cloud Formation: Statistical Insights from LES Cases (model output).* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.8241941` doi: 10.5281/zenodo.8241941

Simon, J. S., Bragg, A. D., Dirmeyer, P. A., & Chaney, N. W. (2021). Semi-coupling of a field-scale resolving land-surface model and wrf-les to investigate the influence of land-surface heterogeneity on cloud development. *Journal of Advances in Modeling Earth Systems*, *13*(10), e2021MS002602.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Duda, M., … Powers, J. (2008). A description of the Advanced Research WRF version 3. *NCAR Technical Note, NCAR/TN-475+STR*.

Taylor, C. M., Gounou, A., Guichard, F., Harris, P. P., Ellis, R. J., Couvreux, F., & De Kauwe, M. (2011). Frequency of Sahelian storm initiation enhanced over mesoscale soil-moisture patterns. *Nature Geoscience*, *4*(7), 430–433.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272.

Weaver, C. P. (2004). Coupling between large-scale atmospheric processes and mesoscale land–atmosphere interactions in the us southern great plains during summer. Part II: Mean impacts of the mesoscale. *Journal of Hydrometeorology*, *5*(6), 1247–1248.