Appendix A Model parameterization for interventions

In this study, the model parameterization follows exactly the same as in Brousse et al. (2023). This means that the model is run at a resolution of 1 x 1 km with 210 by 180 horizontal grid points and 71 vertical layers nested in two other domains run at 12 and 3 km, respectively, and forced by ERA5 6-hourly data at 25 km horizontal resolution. We supplement the version 4.3 to the more recent 4.4.1 to benefit for certain bug fixes; these did not relate to the urban scheme. Of importance, sea surface temperatures are updated every 6 hours out of ERA5, no lake models are activated for inland water bodies, and initial land surface conditions are provided by the default MODIS 5-arc-second land use data set. The latter are further interpolated by the Noah-MP land surface model (Niu et al., 2011) in its default parameterization over 4 soil layers. Building effects on the local climate are calculated by activating the BEP-BEM models (Martilli et al., 2002; Salamanca et al., 2010; Salamanca & Martilli, 2010) which calculates energy fluxes within the urban tile of each grid points – given by the urban fraction. Changes in urban vegetation like in VG_100 are therefore only treated by Noah-MP within the natural tile of grid points where the urban fraction is not null. Fluxes are then averaged at the bulk level to estimate prognostic variables at the each grid point (e.g., air temperature). We chose to use the common Bougeault-Lacarrère planetary boundary layer scheme for these simulations (Bougeault & Lacarrere, 1989). More information on the physical parameterization and on the buildings' thermal and radiative properties can be found in Brousse et al. (2023) under the section 2.a.

The impact of the air conditioning in BEP-BEM is estimated by means of a simple Building Energy Module (BEM). This module computes an energy budget of the indoor air by considering the heat generated by people and equipment, the diffusion of heat through walls and roof, the air infiltration/ventilation, and the radiation entering through the windows. When the indoor air temperature reaches maximum value fixed by the user, the internal temperature is kept constant, and all the extra heat (Hsneed) is ejected to the atmosphere. In addition to the atmosphere is added also the heat generated by the A.C. equipments to do the work ($Hsneed/COP$, where COP is the Coefficient of performance of the A.C. system). In the same way the energy consumption due to the A.C: is estimated as $Hsneed/COP$. In our AC_100 simulations, all buildings are equipped with AC systems running with a COP of 3.5 and a target temperature of 294.15 K. More details can be found in Salamanca et al. (2010); Salamanca and Martilli (2010).

The roof mitigation strategies parameterizations are based on the work of Zonato et al. (2021). The land-surface scheme for green roofs has been developed based on De Munck et al. (2013). It calculates energy and water budgets, taking into account incoming net radiation, water input from precipitation and irrigation (the latter considered as irrigation at the green roof surface), evapotranspiration from vegetation, heat exchange with the atmosphere, and diffusion of energy and moisture throughout the soil. A green roof consists of 10 layers with a total depth of 0.3 m, 5 of them represents the vegetation and the soil substrates, while the rest the underlying roof, including the waterproof membrane. The kind of vegetation present in the upper level is parameterized depending on leaf area index and maximum stomatal resistance.

The parameterization taking into account the effects of RPVPs assumes the photovoltaic arrays to be parallel and detached from roofs and composed of a single layer, and it is based on the work of Jones and Underwood (2001). A photovoltaic panel is assumed to be detached from the roof, and to composed of three layers, as in: a monocrystalline silicon PV cell, a polyester trilaminate and a glass face, with a total depth of 6.55 mm. The prognostic equation of its temperature, that is necessary for calculating the incoming/outgoing heat fluxes, considers: 1) The net incoming short- and long-wave radiation at both surfaces of the photovoltaic panel, assuming a view factor for the bottom surface depending on the area covered by the photovoltaic panel and on its distance to the underlying roof; 2) The heat fluxes dependent on wind speed and temperature differences between the panel and the air (Scherba et al., 2011); 3) The energy produced by the PV cell, dependent on its efficiency and on the PV temperature itself.

The standard WRF version 4.4.1 has been appropriately modified in order to consider a grid-specific ratio of green roof or photovoltaic panels independent of the LCZ, thus independent of look-up tables.

Appendix B Observational Data and Model Evaluation

The model evaluation was performed following the strategy described in Brousse et al. (2023) and uses the same data set. Briefly put, personal weather station air temperature measurements from the *Netatmo* company are gathered using an open API. Each personal weather station's data undergoes a statistical quality-check to ensure that the quality of the measurement is sufficient to perform urban climate studies and model evaluation (Napoly et al., 2018). These crowd-sourced measurements complement the official weather stations measurements coming from the UK MetOffice MIDAS network (UKMO, 2021). Hammerberg et al. (2018) and Brousse et al. (2023) indeed demonstrated that crowd-sourced weather data are beneficial for evaluating urban climate simulations and we therefore decided to take advantage of them in this study too. More information on the data gathering and treatment can be found in Brousse et al. (2023) and related codes can be found following `https://github.com/oscarbrousse/JAMC_BiasCorrection_PWS/`.

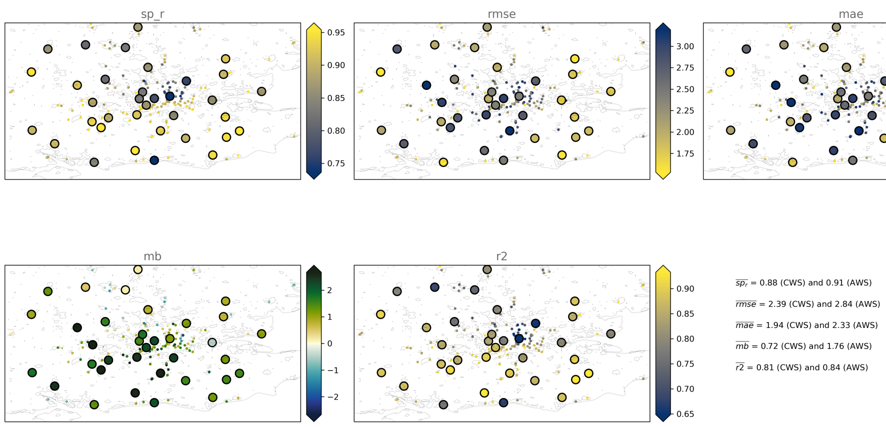

Figure B1. Model evaluation against Netatmo personal citizen weather stations (CWS, small dots) and MIDAS official automatic weather stations (AWS, big dots). Yellow is better. For MB white is the better. Average scores amongst all stations are given in the bottom right.

669 # Appendix C Additional Figures

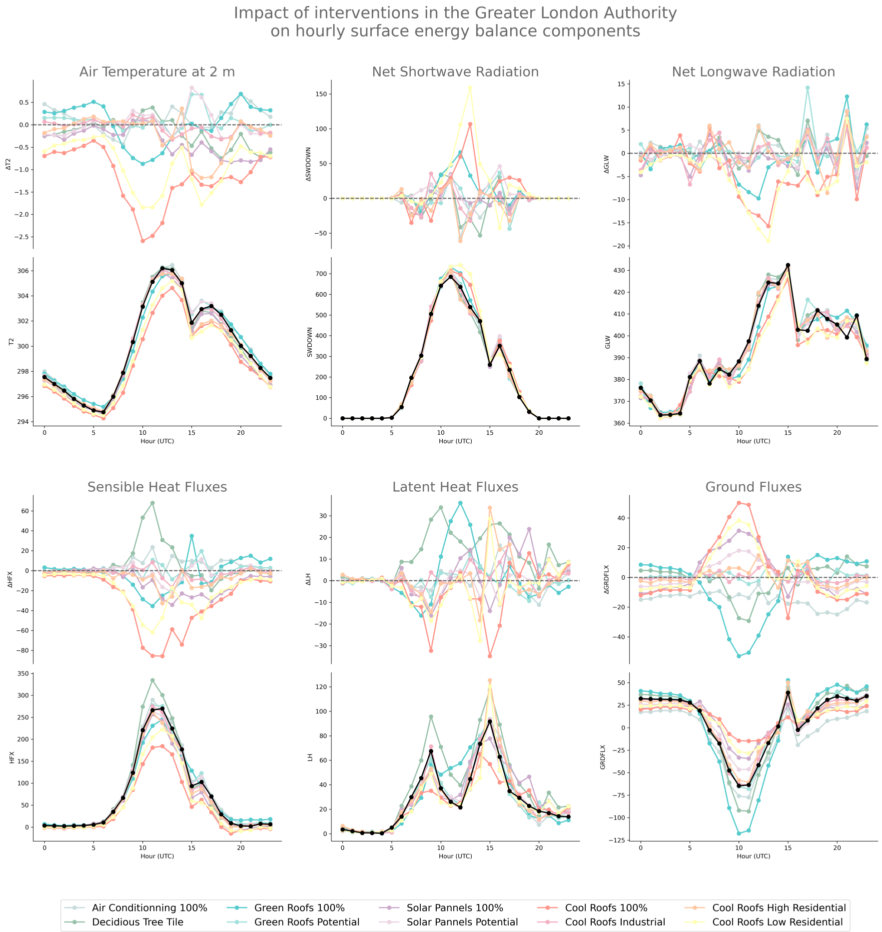

Figure C1. Diurnal cycles of spatially average air temperature at 2 m and of each surface energy balance component (lower panels) and their respective differences (control - intervention; upper panels) over GLA . The black solid line is the control run, the interventions' colors are given in the legend. The dashed grey line represents a null change between the control and the intervention run.

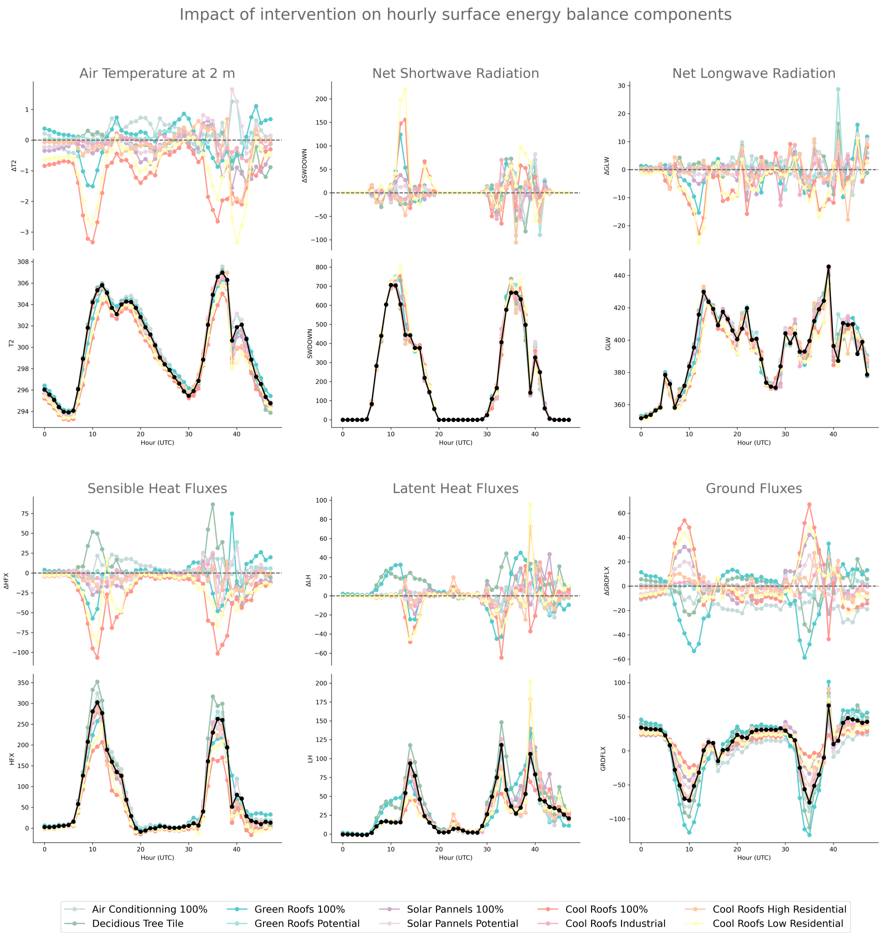

Figure C2. Same as Fig. C1 but without hourly averaging (for the two full days)

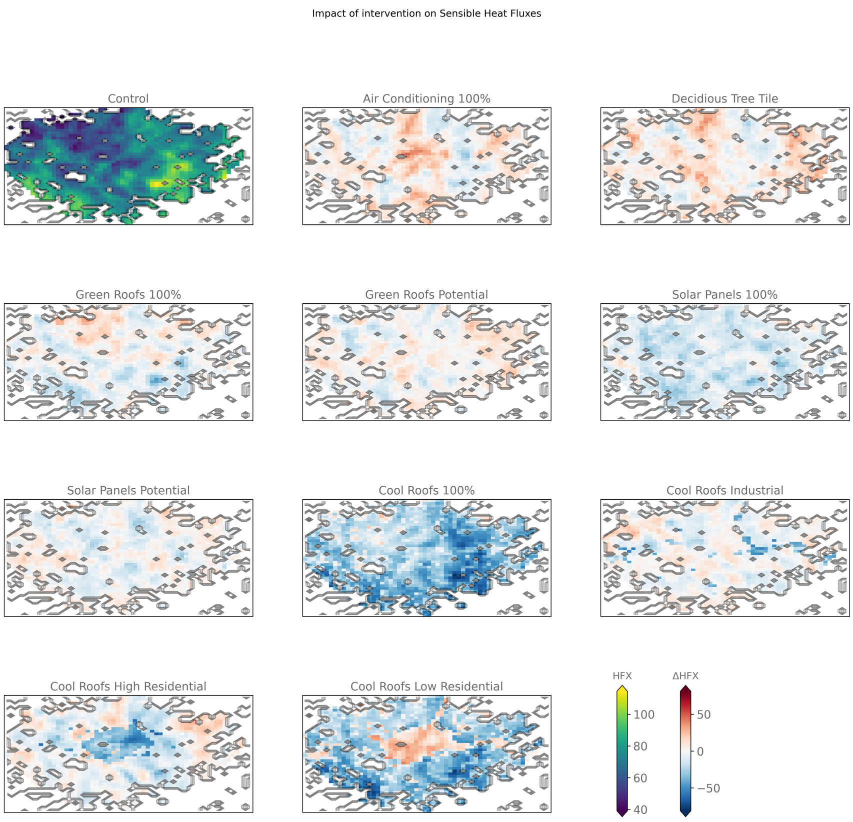

Figure C3. Same as Fig. 1 but for sensible heat fluxes

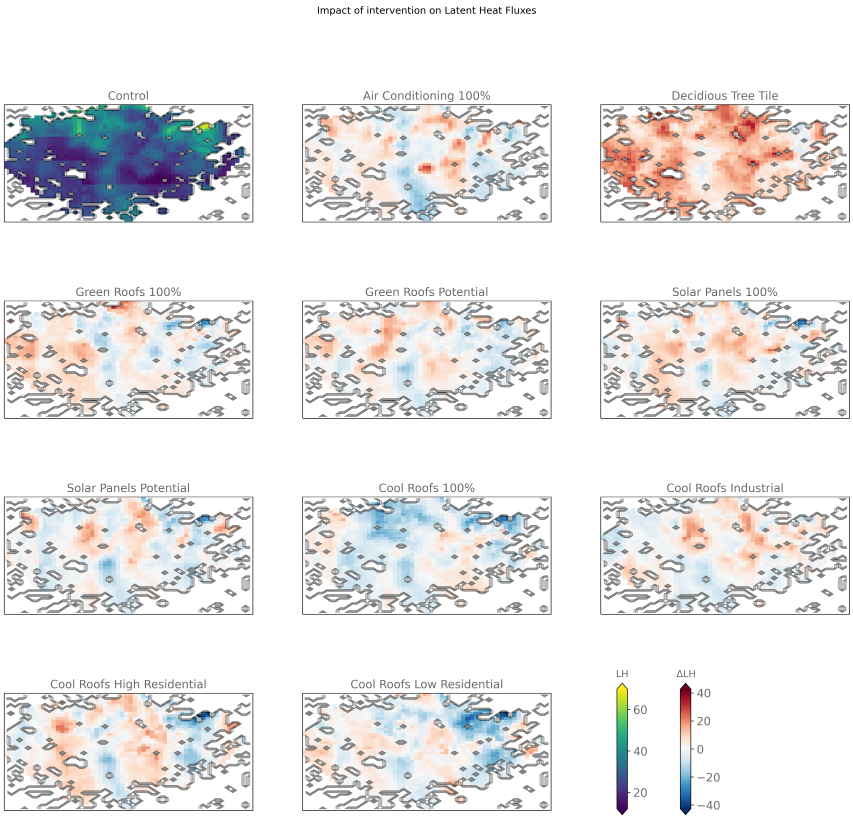

Figure C4. Same as Fig. 1 but for latent heat fluxes

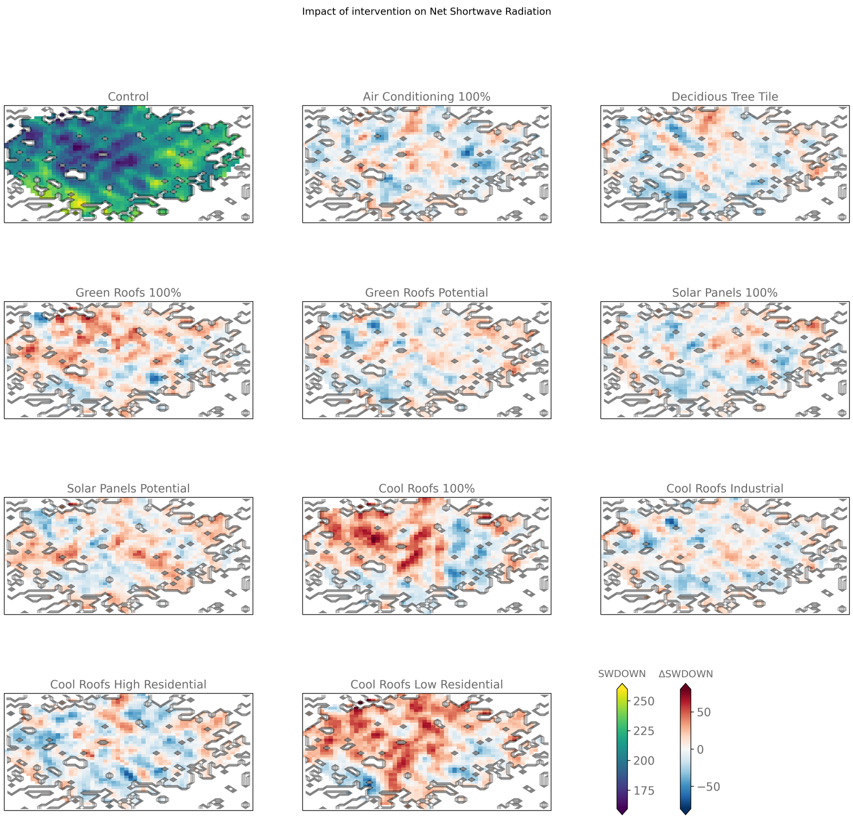

Figure C5. Same as Fig. 1 but for net incoming short-wave solar radiation

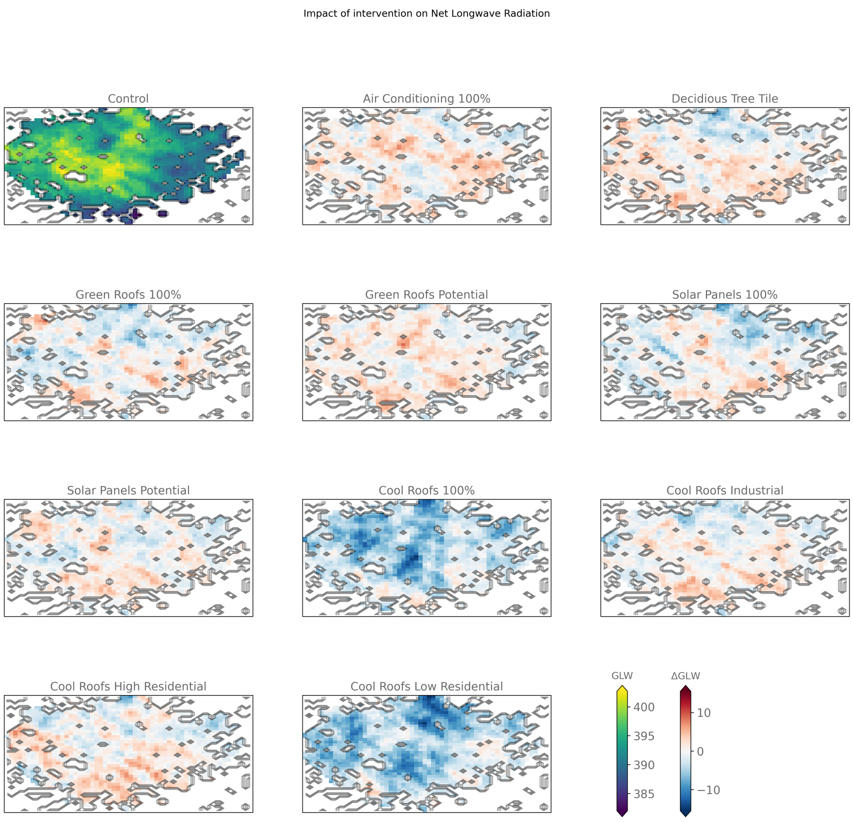

Figure C6. Same as Fig. 1 but for net incoming long-wave radiation

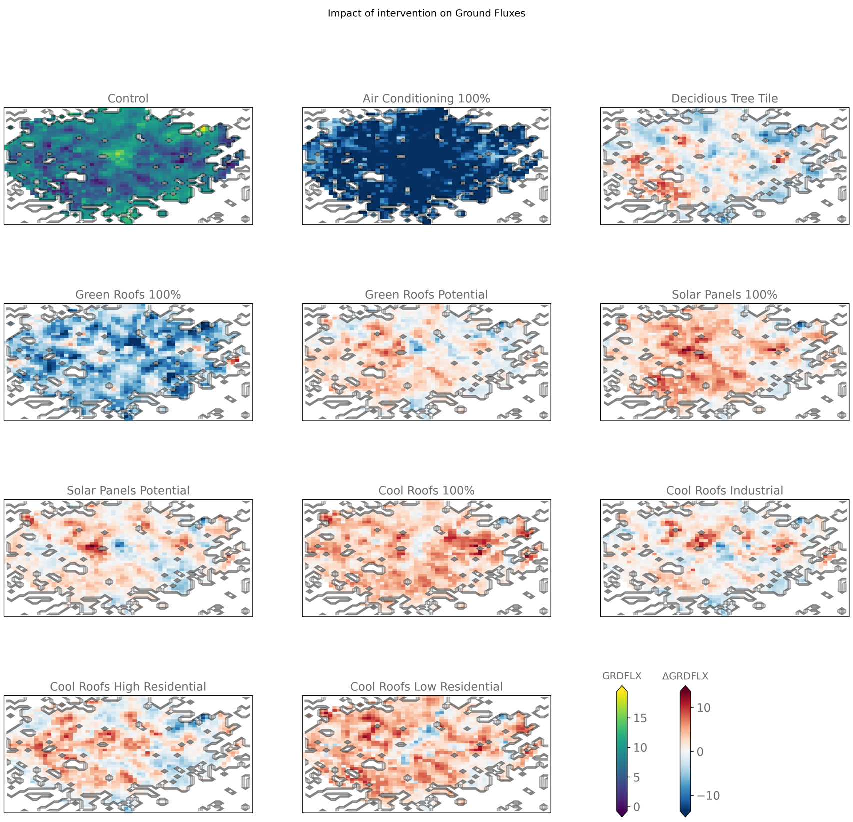

Figure C7. Same as Fig. 1 but for ground fluxes

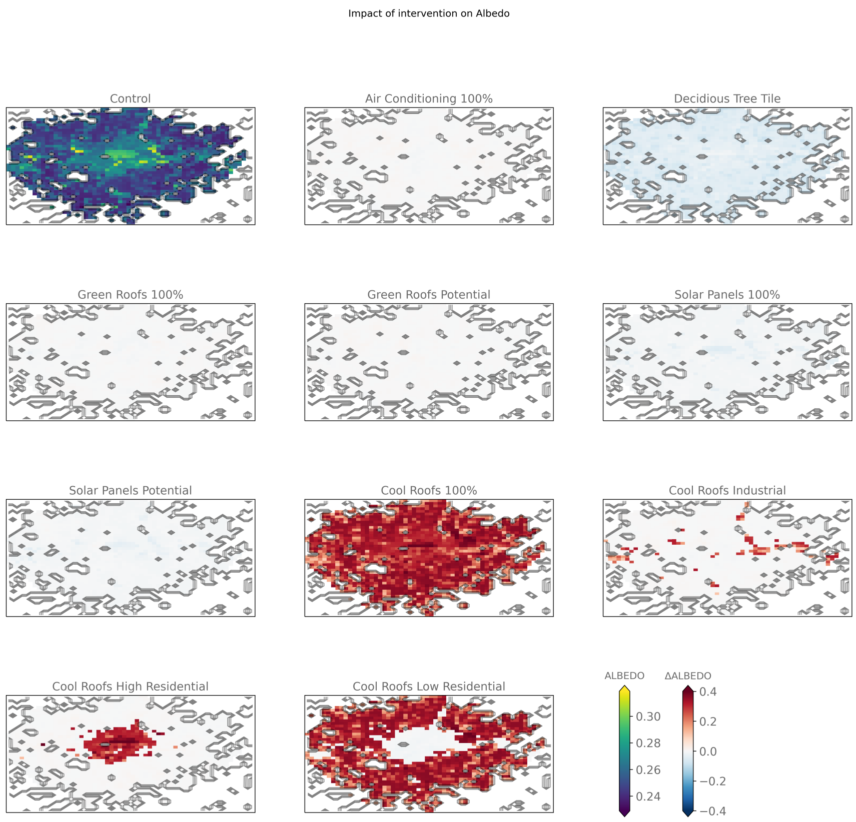

Figure C8. Same as Fig. 1 but for albedo