We deal with the numerical solution of time-dependent partial differential equations by the space-time discontinuous Galerkin method which admits the use of varying meshes at different time levels in a simply way. We develop anisotropic hp-mesh adaptive method which generates anisotropic triangular grids with varying polynomial approximation degrees using the continuous mesh and error models. Whereas the shape of elements and polynomial approximation degrees are optimized locally, the size of mesh elements is optimized globally. The idea is to control the interpolation error in the $L^{\infty}(0,T; L^q(\Omega))$-norm. We present the theoretical as well as practical aspects of this method and demonstrate its efficiency by several numerical examples.