

Average reward over episode using baseline

