1

a) It did perform better in the first phase, because Dyna-Q+ encourages the agent to explore previously unvisited states and therefore finding the goal state faster.
In the second phase the Dyna-Q+ algorithm finds the shortcut faster because it did not visit the neighbouring states for a long time and therefore tries them again.

b) The model needs to be able to handle stochastic processes. But besides this minor fix, the algorithm on slide 13 also works In a stochastic environment.