Observations from the experiments done on different cases used in Windy Gridworld task

My simulator is GUI based. It is very simple to input different data. Users can even change the Start Position, Goal Position, size of the grid window, alpha, gamma, epsilon and most importantly wind intensity. There are options to switch from Standard-Move case to King's Move case. One can enable or disable the wind stochasticity. After input, if the submit button is pressed, a windy gridworld window appears. In the window, one can well analyze the best possible path (episode) that the running experiment could get after long run. Two plots, one shows the convergence of the episode length and other, number of episodes along the time axis, get saved (automatically) in the current running directory.

Given the situations (in the Example 6.5), such as intensity of the wind, alpha, gamma, epsilon, the following observations are:

case1:
alpha= .1, gamma=1, epsilon =.1, wind= [0 0 0 1 1 1 2 2 1 0] … (Described in the fig 6.10)
Grid Size : column =10, row =7, Start = (3,0) Goal =(3,7)   ….. (Coordinate system with (0,0) as origin)

1. Standard Move (Example 6.5):
   - As it has limited moves than king's move, It has slow convergence rate to optimality.
   - It takes too much time (on average more than 1200 time steps) to hit to goal in the first time. i.e length of the first episode is very large.
   - Even after 15000 time steps, it remains far from optimality.
   - Total number of episodes in 15000 time steps is not more than 180
   - Length of the episodes will be min (15) if the steps size to huge
   - As the number of moves is less, it takes more time than Kings'move case to reach to goal.
   - After 600 episode, the episode length started settle down toward min value (15).

2. King's Move (Exercise 6.6):
   - Though it has more move (at max 8) than the standard move (at max 4), it has slow convergence rate but better than Standard move case.
   - It also takes too much time (on average more than 1700 time steps), even worser than Standard Move case, to hit to the goal in the first. i.e. Length of the first episode is very large.
   - Minimum length of the episodes is observed as 7
   - One interesting observation is that, there are more than one best known path with equal length (7).

- Total number of episodes in 15000 time steps is not more than 180.
- After 300 episode, the episode length started settle down toward min value (7).
- As the number of moves is more, it takes less time than Standard moves case.
-

3. Stochastic wind with king's move (Exercise 6.7):
- This experiment is done with king's under stochastic wind environment. The result is highly fluctuating. it always try to converge to best result (average 7 episode length), but does not settle down.
- There are more than 1000 number of episodes in 15000.
- Most interesting with this experiment is that, randomized case is performing better even though it is unable to learn properly.
- It does not take much time to hit to goal in the first time. But length of the first episode on average can not be said perfectly. But in general hits to goal faster than previous cases.
- One interesting observation is that, there are more than one best known path with equal length (7).