***FIRST MODEL : MODEL-BASED, Q VALUE ALGO***

***1 - What is the space of states X?***

The space of states consists of all tuples :

(naive approx)

#states =10\*10\*10\*10\*10\*24 = 2 400 000 states ..

Where is the current charge of our battery set and is the quantity of H2 currently stored in our vessel, and is the current price of electricity (fluctuating one). T is the current time of the day

* 5 continuous quantities that will be discretized
* The transitions can be estimated from data (all state variables are deterministic, juste needs to be estimated from data

***2 - What is the space of actions A?***

The space of actions consists of all tuples :

MW ( ???)

Once again, all these flows are continuous quantities (although we can consider a discretization).

* Both space of states and actions are continuous. Not ideal

***3 - Can we discretize A and X ?***

*As such, it is absolutely impossible to discretize the state of actions : even by considering a very coarse discretization, let’s say we end up with*  …

An alternative would be to reduce the dimension of by considering the Solar and Wind farm end up being one unique source of energy, and that actions have to be taken on this source only instead of the two farms separately. That would make us save 3 dimensions. If we also assume that H2 produced by the electrolyzer must 1st be stored before being distributed, we can save another 3 dimensions, and we end up with : :

We can again reduce the numbers of dimensions by using the fact that all flows originating from energy sources must sum to 100% (no waste). So again, we can save 1 other dimension.

Now if we go back to our initial discretization, we only have , which is fairly low.

We could refine the discretization: 5 🡪 10 possible values for each flow, and end up with 1 million actions possible.

***3 – What are the observations?***

We can consider that our observations at every step consist of :

1. The current purchase/retail price of electricity (do we model that ? )
2. The amount of electricity produced by our Solar farm
3. The amount of electricity produced by our Wind farm

***4 – Is our model observable?***

Yes, we assume we can access our state variables at all time.

***5 – Do we know our transition function?***

In the simplest case, our transitions are fully deterministic : given a given state, taking a given action will unambiguously lead us to a unique state.

If we want more realistic scenarios, we must add some white noise to the final values

***6 – Do we know our reward function?***

The reward function is also fully given at a specific timestep by the current price of electricity + action taken : it is simply given by the collected money. We may also want to penalize some undesirable events, in which case the reward function will have to adapt.