

To investigate the reason why the number of patches can affect the number of parameters in the model. We use Figures S1–3 as examples to explain.

```

//////////Mamba//////////
Mamba (
  d_model=mlp_hidden_dim//2,
  d_state=4, # SSM state expansion factor
  d_conv=4, # Local convolution width
  expand=2, # Block expansion factor
)

```

Figure S1: Represents the mamba layer's parameter we set in the models

```

//////////Mamba block in Mucm-Net(1-patch)//////////
(fc21): MambaLayer (
  (mamba): Mamba (
    (in_proj): Linear(in_features=8, out_features=32, bias=False)
    (conv1d): Conv1d(16, 16, kernel_size=(4,), stride=(1,), padding=(3,), groups=16)
    (act): SiLU()
    (x_proj): Linear(in_features=16, out_features=9, bias=False)
    (dt_proj): Linear(in_features=1, out_features=16, bias=True)
    (out_proj): Linear(in_features=16, out_features=8, bias=False)
  )
)

```

Figure S2: Snapshots on one of the Mamba layers in MUCM-Net (1-patch)

```

//////////Mamba block in Mucm-Net(2-patch)//////////
(fc21): MambaLayer (
  (mamba): Mamba (
    (in_proj): Linear(in_features=4, out_features=16, bias=False)
    (conv1d): Conv1d(8, 8, kernel_size=(4,), stride=(1,), padding=(3,), groups=8)
    (act): SiLU()
    (x_proj): Linear(in_features=8, out_features=9, bias=False)
    (dt_proj): Linear(in_features=1, out_features=8, bias=True)
    (out_proj): Linear(in_features=8, out_features=4, bias=False)
  )
)
(fc22): MambaLayer (
  (mamba): Mamba (
    (in_proj): Linear(in_features=4, out_features=16, bias=False)
    (conv1d): Conv1d(kernel_size=(4,), stride=(1,), padding=(3,), groups=8)
    (act): SiLU()
    (x_proj): Linear(in_features=8, out_features=9, bias=False)
    (dt_proj): Linear(in_features=1, out_features=8, bias=True)
    (out_proj): Linear(in_features=8, out_features=4, bias=False)
  )
)

```

Figure S3: Snapshots on one of the Mamba layers in MUCM-Net (2-patch)

From Figures S2 and S3, we can calculate the number of parameters in those layers. The number of parameters for the layer sample in MUCM-Net with 1 patch is 640. However, the number of parameters for the layer sample in MUCM-Net with 2 patches is 448, which is the sum of two Mamba layers for the same input and output shape. As shown in Figures S2 and S3, as the number of patches increases, the number of parameters decreases.