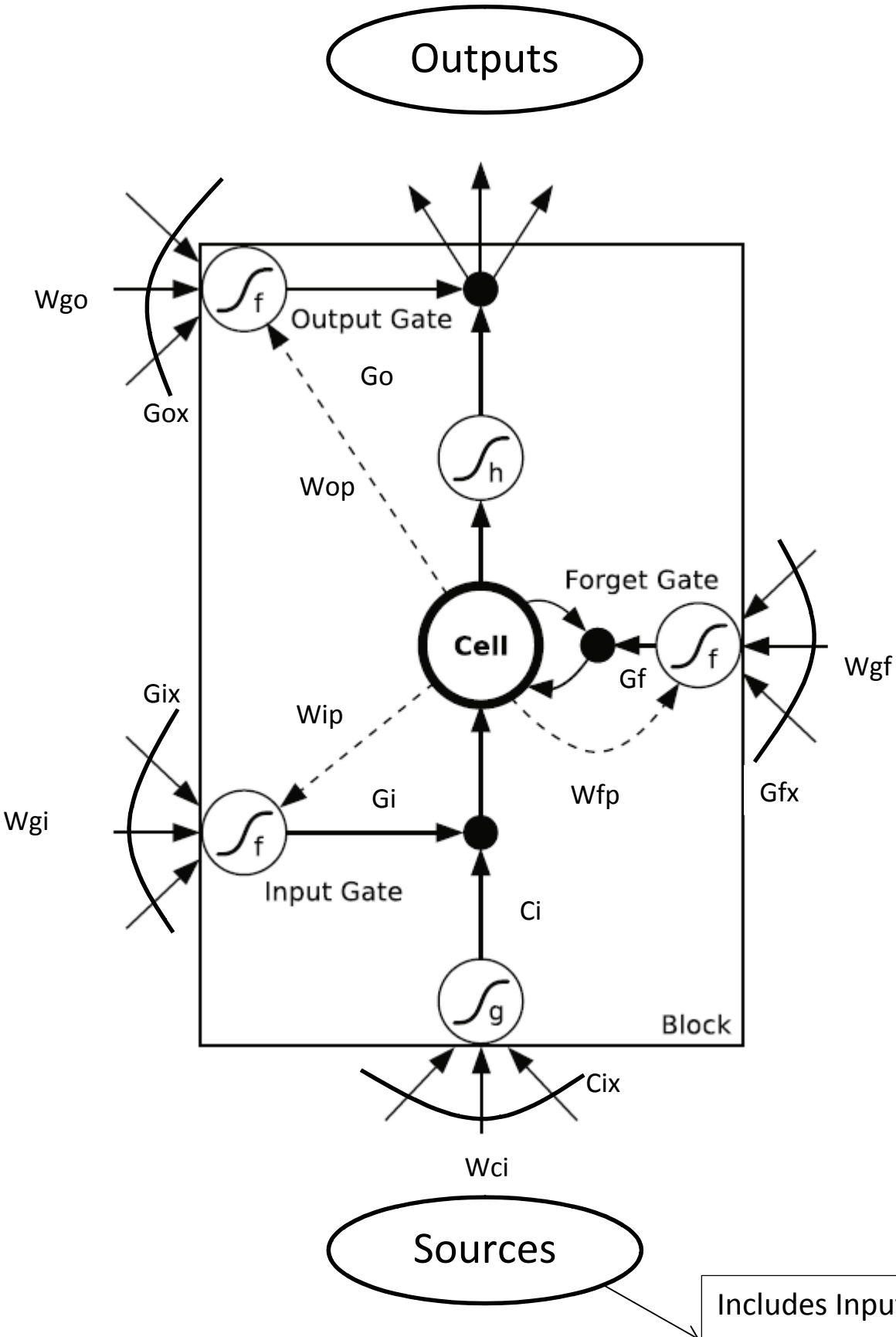

Includes Input features at time t and outputs from t-1

Comparing lstm.py and the equations in Supervised Sequence Labeling using Recurrent Neural Networks:

- Wgi = $w_{i\iota}$ and $w_{h\iota}$ augmented
- Wgo = $w_{i\omega}$ and $w_{h\omega}$ augmented
- Wgf = $w_{i\phi}$ and $w_{h\phi}$ augmented
- Wci = w_{ic} and w_{hc} augmented
- Wip = $w_{c\iota}$
- Wop = $w_{c\omega}$
- Wfp = $w_{c\phi}$
- Source[t] = x^t and b_h^{t-1}
- State[t] = S_c^t
- Output[t] = b_c^t
- Cix[t] = a_c^t
- Ci[t] = $g(a_c^t)$
- Gix[t] = a_ι^t
- Gi[t] = b_ι^t
- Gox[t] = a_ω^t
- Go[t] = b_ω^t
- Gfx[t] = a_ϕ^t
- Gf[t] = b_ϕ^t
- outerr[t] = $\square_c^t$
- stateerr[t] = $\square_s^t$

- gierr[t] =δ_{ι}^{t}
- gferr[t] = δ_{ϕ}^{t}
- goerr[t] = δ_{ω}^{t}
- cierr[t] = δ_{c}^{t}