Appendix A Metapopulation Models A - D

The metapopulation Models, A and B, have the Ricker map as local dynamics with the sequence $\mathbb{NDR}$. The full dynamical equation for each subpopulation i at time t is,

$$X_{i,t+1} \equiv (\mathbb{NDR}\mathbf{X}_t)_i = \left\{(1-\epsilon)X_{i,t}\exp[r(1-X_{i,t})] + \frac{\epsilon}{4}\sum_{\langle j;i\rangle} X_{j,t}\exp[r(1-X_{j,t})]\right\}\exp(\lambda\zeta_{i,t}). \quad \text{(A1)}$$

The growth parameter r for Model A is 2.2 and for Model B is 2.4. Both growth parameters are in the two-cycle region.

Metapopulation Model C has the logistic map as local dynamics with the sequence $\mathbb{NDR}$. The full dynamical equation is,

$$X_{i,t+1} \equiv (\mathbb{NDR}\mathbf{X}_t)_i = \left\{(1-\epsilon)rX_{i,t}(1-X_{i,t}) + \frac{\epsilon}{4}\sum_{\langle j;i\rangle} rX_{j,t}(1-X_{j,t})\right\}\exp(\lambda\zeta_{i,t}) \qquad \text{(A2)}$$

where the growth parameter $r = 3.2$.

Metapopulation Model D has the Ricker map as local dynamics with growth parameter $r = 2.2$ but with a different sequence, $\mathbb{NRD}$ from Modal A. The full dynamical equation is,

$$X_{i,t+1} \equiv (\mathbb{NRD}\mathbf{X}_t)_i = \left\{[(1-\epsilon)X_{i,t} + \frac{\epsilon}{4}\sum_{\langle j;i\rangle} X_{j,t}]\exp\left[r(1-[(1-\epsilon)X_{i,t} + \frac{\epsilon}{4}\sum_{\langle j;i\rangle} X_{j,t}])\right]\right\}\exp(\lambda\zeta_{i,t}). \quad \text{(A3)}$$

In addition to the choice of growth parameter r of the local dynamics and the sequence of processes, each metapopulation model has three other parameters: the number of patches, N, dispersal, ϵ, and noise, λ.

The state diagram separating synchronous and incoherent regions in dispersal and noise plane $(\epsilon - \lambda)$ for metapopulation Models A-D is shown in Fig. A1.

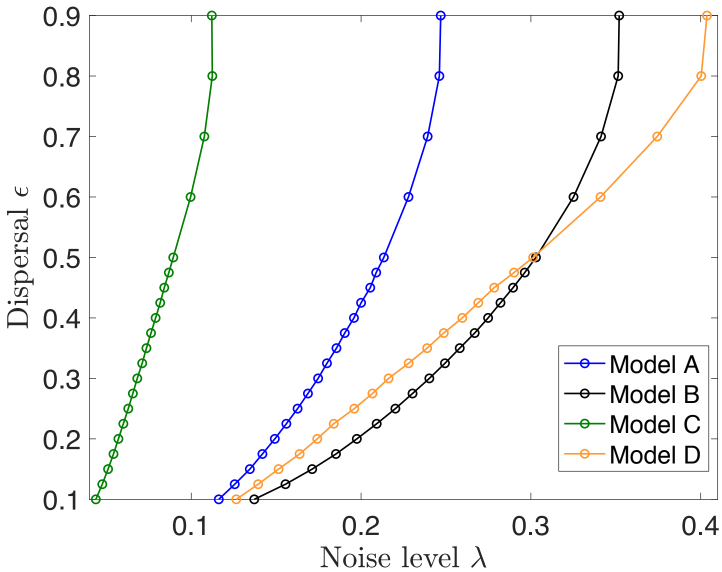

Figure A1: State diagram separating synchronous (low noise and high dispersal) and incoherent regions (high noise and low dispersal) for the metapopulation Models A-D studied here.

Appendix B Binder cumulant method

The state diagram for the metapopulation Models A-D is plotted by finding the critical noise $\lambda_c(\epsilon)$ for various dispersal values, ϵ using the Binder cumulant method. The Binder cumulant

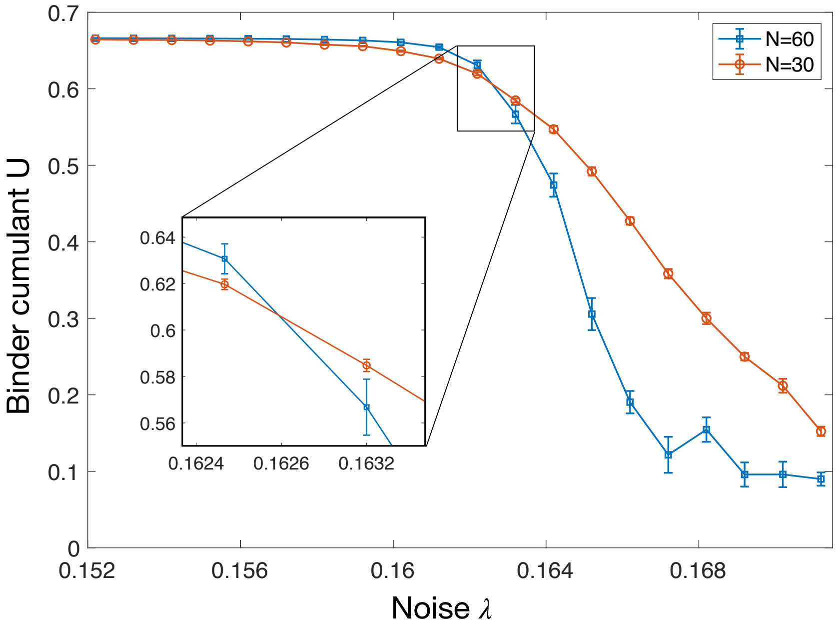

Figure A2: The critical noise $\lambda_c(\epsilon)$ for dispersal $\epsilon = 0.25$ is calculated for metapopulation Model A as the crossings of the Binder cumulant curves for patch sizes $N = 30, 60$. The inset shows the crossings of the Binder cumulant curves at $\lambda_c(\epsilon) = 0.1626$.

is a fourth-order cumulant of synchronization variable,

$$U = 1 - \frac{\langle \tilde{s}_t^4 \rangle}{3\langle \tilde{s}_t^2 \rangle^2} \tag{A4}$$

where $\tilde{s}_t = \frac{1}{N}\sum_i \tilde{S}_{i,t}$ is the instantaneous synchronization variable and $\langle . \rangle$ computes the long-time average. In the synchronous region, the Binder cumulant takes the value 2/3 whereas it takes the value zero in incoherent region. At the critical transition, the critical Binder cumulant on a 2D lattice with periodic boundary conditions takes the value $U^* = 0.6169$ [5].

The critical noise $\lambda_c(\epsilon)$ for a given dispersal ϵ is obtained from the crossings of the Binder cumulant curves plotted as function of noise for different number of patches, N. In this work, we use $N = 30, 60$ to find the critical noise from the crossings of the Binder cumulant curve. The Binder cumulant curves are shown for dispersal $\epsilon = 0.25$ as function of noise λ in Fig. A2. The Binder cumulant curves crosses at critical noise $\lambda_c(\epsilon) = 0.1626$ as shown in the inset.

Appendix C Dynamical Ising model with memory

The dynamical Ising model with memory [4, 1] has nearest neighbor coupling J and a self-interaction strength K. The single spin transition probability, $P(S_{i,t+1}|\mathbf{S}_t, J, K)$, is given by,

$$P(S_{i,t+1}|\mathbf{S}_t, J, K) = \frac{\exp\left[(Jh_{i,t} + KS_{i,t})S_{i,t+1}\right]}{2\cosh\left[Jh_{i,t} + KS_{i,t}\right]} \tag{A5}$$

where the local field, $h_{i,t}$, is the sum of the nearest neighbors of $S_{i,t}$ at time t

$$h_{i,t} = \sum_{\langle j;i \rangle} S_{j,t}. \tag{A6}$$

The transition probability for the full spin configuration for parallel dynamics (in which all

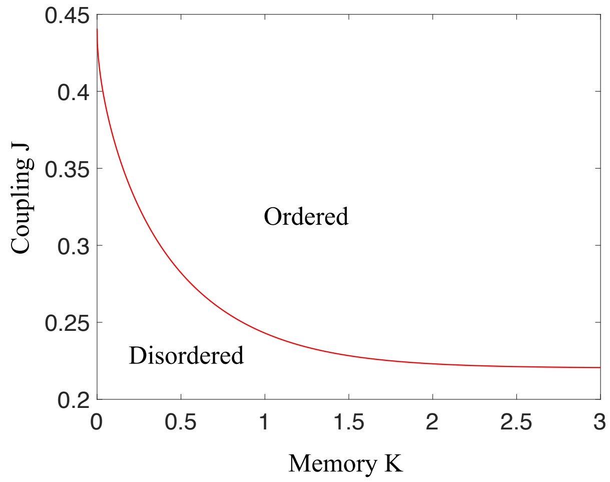

Figure A3: State diagram of the dynamical Ising model with memory K and nearest neighbor coupling J [1]. The solid red line is the critical line separating the ordered (synchronous) and disordered (incoherent) regions in the J–K plane. The tendency toward ordering increases with both J and K. The point $K = 0$, $J_c(0) = 0.44\ldots$ is the critical point of the standard Ising model. At large K, the critical coupling approaches half the value of the standard Ising model.

spins are updated simultaneously each time step) is given by,

$$P(\mathbf{S}_{t+1}|\mathbf{S}_t, J, K) = \prod_i P(S_{i,t+1}|\mathbf{S}_t, J, K) = \prod_i \frac{\exp[(Jh_{i,t} + KS_{i,t})S_{i,t+1}]}{2\cosh[Jh_{i,t} + KS_{i,t}]}. \tag{A7}$$

Note that the memory (K) and nearest neighbor coupling (J) parameters are both dimensionless. It is also common to write these parameters as ratios, J/T and K/T where now J and K are considered interaction "energies" and T is a "temperature" and reflects the noise level in the system. Because only the ratios appear in Eq. (A7), it is not possible to distinguish the difference between increasing noise and decreasing interaction strength, without making additional assumptions.

The transition probability for the standard dynamical Ising model is obtained by setting $K = 0$ so the transition probability for spin, $S_{i,t+1}$ depends only on its neighbors and not on itself. This is in contrast to ecological dynamics, in which the current local state is a density dependent function of the previous one.

The Ising model with memory undergoes a critical transition in the Ising universality class from a disordered to an ordered state. For a fixed value of K this transition occurs at a value, $J_c(K)$. The critical line, $J_c(K)$, separating ordered and disordered states in the coupling-memory (J–K) plane is shown in Fig. A3 [1].

Appendix D Inference methods

In this section we introduce maximum likelihood inference methods to determine the parameter values $\{J, K\}$ for the dynamical Ising model with memory that best describe the dynamics of the metapopulation model. Inferring an Ising model from data is called an inverse Ising problem [6]. Inverse Ising problems typically refer to inferring parameters of the stationary Gibbs distribution that are most like the data sampled at a single time. By contrast, here we use *dynamical* inference [2] to find the dynamical Ising model that best represents the dynamics of the metapopulation model. Of course, if the dynamical model is accurate, its stationary states should be close to the stationary states of the metapopulation.

The likelihood that we maximize is related via Bayes' theorem and several assumptions to the transition probability of the Ising model, $P(\mathbf{S}_{t+1}|\mathbf{S}_t, J, K)$ (see Eq. (A7)).

We are interested in the likelihood $L\big(\{J,K\}|\tilde{\mathbf{S}}_{t+1}, \tilde{\mathbf{S}}_t\big)$ of the parameter values {J,K} given spin configurations from two successive time steps, $\{\tilde{\mathbf{S}}_{t+1}, \tilde{\mathbf{S}}_t\}$ of the metapopulation. Using Bayes' theorem and the assumption of uniform priors for both the parameters and the spin configurations, the likelihood is proportional to the transition probability of the Ising model with memory, $P(\mathbf{S}_{t+1}|\mathbf{S}_t, J, K)$ (Eq. A7). It is equivalent but more convenient to maximize the logarithm of the likelihood, $\mathcal{L} = \log L$,

$$\mathcal{L}\big(\{J,K\}|\tilde{\mathbf{S}}_{t+1}, \tilde{\mathbf{S}}_t\big) \propto \sum_i \Big\{[J\tilde{h}_{i,t} + K\tilde{S}_{i,t}]\tilde{S}_{i,t+1} - \log[2\cosh(J\tilde{h}_{i,t}S_{i,t} + \tilde{K})]\Big\}. \tag{A8}$$

Finally, while this is an exact formula for inferring parameters from data, we assume that maximizing $\mathcal{L}$ using the Ising representation of the metapopulation data $\tilde{\mathbf{S}}_t$ will yield a good Ising model representation. Thus, the goal is to maximize $\mathcal{L}\big(\{J,K\}|\tilde{\mathbf{S}}_{t+1}, \tilde{\mathbf{S}}_t\big)$. Because the Ising spin representation of the metapopulation involves only binary variables and because of additional symmetries of the dynamical Ising model, maximizing the likelihood reduces to a simple numerical fit of the two parameters with respect to ten numbers obtained directly from the data.

Possible values of the product $h_{i,t}S_{i,t}$ for all different spin initial values at time t are $\{-4,-2,0,2,4\}$ and for $S_{i,t}S_{i,t+1}$ are $\{-1,1\}$. Considering $\mathcal{Z}_2$ symmetry of the model, there are only 10 different possible values for each term in the sum in Eq. (A8). Also, the sum of probabilities (Eq. A7) with the two possible values of $S_{i,t}S_{i,t+1}$ with same $h_{i,t}S_{i,t}$ must be one. This can be understood as given the product $h_{i,t}S_{i,t}$, spin $S_{i,t+1}$ either flips or not which is equivalent as $S_{i,t}S_{i,t+1}$ being either -1 or 1.

The flip probability given the product is $P_f(h_{i,t}S_{i,t})$ as shown in Eq. (8). Whenever $S_{i,t}S_{i,t+1} = -1$, the corresponding probability is $P_f(h_{i,t}S_{i,t})$ and when $S_{i,t}S_{i,t+1} = 1$ the probability is $1 - P_f(h_{i,t}S_{i,t})$.

Now, spins in $\tilde{\mathbf{S}}_t$ can be grouped into 5 bins corresponding to their products $h_{i,t}S_{i,t}$. Lets say the number of spins in each bin is n_k where k represent the values of the product $h_{i,t}S_{i,t}$ $(\{-4,-2,0,2,4,\})$ then $\sum_k n_k = N$ with N being the total number of spins. Out of these n_k spins in a bin, n_k^f spins chose to flip at time $t+1$ with $S_{i,t}S_{i,t+1} = -1$ which enforces $n_k - n_k^f$ spins with no flip at $t+1$. Then the above log-likelihood function in Eq. (A8) can be rewritten as,

$$\mathcal{L}\big(\{J,K\}|\tilde{\mathbf{S}}_{t+1}, \tilde{\mathbf{S}}_t\big) = \sum_k \Big\{n_k^f \log(P_f(k)) + [n_k - n_k^f]\log(1 - P_f(k))\Big\} \tag{A9}$$

Here $\{P_f(k)\}$ are functions of parameters $\{J,K\}$ and n_ks are specific to the spin configurations $\tilde{\mathbf{S}}_t, \tilde{\mathbf{S}}_{t+1}$. So given the spin configurations at time t and $t+1$, the log-likelihood is a function of $\{J,K\}$ which is maximized to infer parameters.

The spin configurations $\tilde{\mathbf{S}}_t, \tilde{\mathbf{S}}_{t+1}$ are obtained over 100 independent runs and the n_ks are collected which form a multinomial distribution. The error bars for the observations are calculated by the bootstrap method [3] on the obtained multinomial distribution. The error bars on the inferred results in Fig. 5 contain only statistical errors. The systematic errors in identifying exact parameters of the metapopulation critical line are not included.

Appendix E Forecast skill

Forecast skill measures the accuracy of a predictive model. We use the Brier skill score to characterize the ability of the inferred dynamical Ising model to predict the metapopulation model one step in the future. The Brier score, BS is the mean squared error in a probabilistic prediction, P_i of a binary observable, F_i averaged over many events k.

$$BS = \frac{1}{M}\sum_{k=1}^{M}(P_k - F_k)^2. \tag{A10}$$

Here F_k is either zero or one and P_k the predicted probability that F_k will be one. For example, in weather forecasting, a binary observable is rain ($F_k = 1$) or no rain ($F_k = 0$) and P_k the probabilistic forecast (e.g. 20% chance of rain).

Brier forecast skill, FS compares the Brier score of the probabilistic prediction to a simple reference forecast,

$$FS = 1 - \frac{BS}{BS_{\text{ref}}} \tag{A11}$$

where BS_{ref} is the Brier score of the reference forecast (e.g. the overall climatological probability of rain). Forecast skill is 1 for a perfect forecast, greater than zero for a forecast that improves on the reference, and negative for forecast with less skill than the reference.

We seek to predict the probability that a subpopulation changes its phase of oscillation ("flips") in the next time step. Given successive two-cycle configurations for a given dispersal and noise, $\mathbf{M}_t$ and $\mathbf{M}_{t+1}$, we obtain the corresponding phase configurations, $\tilde{\mathbf{S}}_t$ and $\tilde{\mathbf{S}}_{t+1}$ (see Eq. (6)). A phase change of subpopulation i occurs at time t if the phase flip variable, $F_{i,t}$,

$$F_{i,t} = \frac{|\tilde{S}_{i,t+1} - \tilde{S}_{i,t}|}{2}, \tag{A12}$$

is one and no phase change occurs if it is zero.

The probabilistic prediction, P_f is made using the flip probability of the inferred Ising model with memory, $P_f = P(S_{i,t+1} = -S_{i,t}|\mathbf{S}_t, J, K)$ (see Eq. (8)) where the parameters J, K are taken from the inference results for the chosen dispersal and noise and $\tilde{\mathbf{S}}_t$ is set by the metapopulation phase variables $\mathbf{S}_t$. The Brier score for the dynamical Ising model with memory is obtained by averaging over the $N = 60^2$ subpopulations of the metapopulation and then over 100 independent samples of the metapopulation in the stationary state. For each subpopulation, the product $\tilde{h}_{i,t}S_{i,t}$ is the input to flip probability. The flip probabilities are shown in Fig. 7, panels (b)-(f). The reference prediction is the overall probability of a flip averaged over the stationary state. This probability is plotted in Fig. 7, panel (a).

Since the metapopulation models are stochastic, the skill scores of the models themselves are less than one. In order to calculate the Brier score of the metapopulation models we use the defining lattice maps, for example Eq. (A1) for Models A and B. For given configurations $\mathbf{X}_t$ and $\mathbf{X}_{t+1}$ we can solve an algebraic equation for a bound on the noise factor $\exp(\lambda\zeta_{i,t})$ for which a flip occurs for a given subpopulation. The probability that this bound is satisfied can be calculated analytically as an error function and is the prediction of the model. The forecast skill score of the model itself is shown as the upper curve in Fig. 8 and is a bound on the performance of any other predictive model of the metapopulation.

Appendix F Simulation details

We studied the dynamics of the four metapopulation models (Table 1) numerically on a square lattice of length 60 ($N = 60^2$) with periodic boundary conditions. For a given metapopulation model, we have carried out simulations for various values of dispersal ($0 < \epsilon < 1$) with noise $\lambda_c(\epsilon)$ chosen such that the parameters are on the critical line (see Fig. 3 for metapopulation model A). Three consecutive lattice configurations were collected for analysis at the end of each run which consists of 3×10^6 timesteps. Results shown in the following sections were obtained from the average over 100 individual runs, corresponding to 100 independent sets of three consecutive metapopulation configurations.

The best fit parameters values $\{J, K\}$ of the dynamical Ising model with memory were obtained using maximum likelihood inference methods (Appendix D). The likelihood is proportional to the transition probability of the Ising model with memory, $P(\mathbf{S}_{t+1}|\mathbf{S}_t, J, K)$ (see Eq. (A7) of Appendix C).

References

[1] V Nareddy and J Machta. Kinetic Ising models with self-interaction: Sequential and parallel updating. *Phys Rev E*, 101:012122, 2020.

[2] Aurélien D, Federico, and Pan Z. Data quality for the inverse Ising problem. *J Phys A Math Theor*, 49(38):384001, 2016.

[3] M Newman and G Barkema. *Monte Carlo methods in statistical physics*. Clarendon Press, Oxford, 1999.

[4] E Cirillo, P Louis, W Ruszel, and C Spitoni. Effect of self-interaction on the phase diagram of a Gibbs-like measure derived by a reversible probabilistic cellular automata. *Chaos Solitons Fractals*, 64:36 – 47, 2014.

[5] W Selke. Critical Binder cumulant of two-dimensional Ising models. *Eur Phys J B*, 51(2):223–228, 2006.

[6] E Aurell and M Ekeberg. Inverse Ising inference using all the data. *Phys Rev Lett*, 108:090201, 2012.