Supporting Information for

Massive and infrequent informed emigration events in a species threatened by climate change: the emperor penguins.

Garnier et al.

Corresponding author: Jimmy Garnier.

E-mail: jimmy.garnier@univ-smb.fr

A Mathematical models and data

[CHECK THAT NOTHING IS MISSING IN THE MODELLING FRAMWORK COMAPRED TO THE MAIN TEXT]

A.1 Genetic data

In our paper, we use the genetic data from Younger *et al.* [2017]. Specifically, we use 4,596 genome-wide single nucleotide polymorphisms (SNPs), characterized in 110 individuals (10 to 16 per colony) from eight colonies around Antarctica (Ammanda Bay, Pointe Geologie, Fold Island, Auster, Cape Roget, Cape Washington, Gould Bay and Halley Bay).

The loci included in the data set have been genotyped by at least 80% of the individuals per population. We can thus assume that the individuals were genotyped at the same loci (with possibly some allele frequencies equal to 0).

In addition, each neutral genetic cluster characterizes a geographic region of Antarctica:

1. Weddell sea (Gould Bay to Halley Bay colonies) (WEDD),
2. Mawson Bay (Fold Island to Cape Darnley colonies) (MAWS),
3. Amanda Bay to Pointe Geologie colonies (AMPG) and
4. Ross sea (Cape Washington and Cape Crozier colonies) (ROSS).

Thus, we assume that before our data collection the individuals located in a colony among the four regions, belong to genetic cluster characterizing the region. To characterize the structure at a circum-Antarctica scale, we assign Davis Bay and Mertz Glacier colonies with the AMPG cluster, and Cape Colbeck and Rupert Coast with the ROSS cluster.

Out of theses four geographic regions, we further include three geographic regions for which no genetic data is available:species

5. from Smith to Snowhill Island in the Wedell sea colonies (StoS),
6. from Stancomb to Kloa point colonies (StoK) and
7. Ledda bay to Rotschild colonies (Admunsen and Bellingshausen seas, A-B seas).

We thus obtain 7 different geographical regions and 4 genetic clusters.

A.2 Demographic model for the emperor penguin

In our analysis, the population dynamics of emperor penguin is described using the meta–population model developed by Jenouvrier *et al.* [2017]. In this section we describe in detail the different function involved in this model, which projects the population vector $\mathbf{n}$—comprising the population size n_i in each colonies i—from time t to $t+1$. We write

$$\mathbf{n}(t+1) = \mathbf{D}\big[t, \mathbf{n}(t)\big]\mathbf{F}\big[t, \mathbf{n}(t)\big]\,\mathbf{n}(t) \tag{1}$$

to indicate that the projection interval is divided into two main phases of possibly different duration: the motionless reproduction phase (**F**) followed by the dispersal phase (**D**). The projection matrices **D** and **F** depend on both the current population density $\mathbf{n}(t)$ and time t because the habitat conditions vary among patches and over time.

The reproduction matrix F

The reproduction matrix **F** is constructed using the Ricker model. It includes the intrinsic population growth rate $\mathbf{r}(t)$ that may vary in time because it depends on sea ice concentrations (SIC), $\mathbf{r}(SIC_t)$. For each projection interval t, we calculate the growth rate of each colony $r_i(t)$ using the median of the stochastic population size projected by a sea-ice dependent population model without density dependence to account for uncertainties related to both climate and demographic processes Jenouvrier *et al.* [2017]. This Ricker model also includes the carrying capacity of the colonies **K** which are assumed to be constant over time.

The dispersal phase D

The dispersal phase (**D**) is decomposed into three stages: (1) emigrating from the resident patch, (2) searching for new patch among other patches with an average dispersal distance d (transfer), and (3) settling in a new patch. During this dispersal event, individuals may select the habitat with highest quality (informed search) or settle in a random habitat (random search). The dispersal projection matrix **D** is thus decomposed as follows

$$\mathbf{D} := \mathbf{S}[t; d]\, \mathbf{M}[t, \mathbf{n}(t); \mathbf{p}_m] \tag{2}$$

to indicate that matrices for searching behavior, **S**, and emigration, **M**, depend on the population size (**n**) as well as the environmental conditions which depend on time t and the coastal distance between the colonies. We contrast three dispersal behaviors: (1) an informed

dispersal behavior where individuals leave poor habitat colony (informed emigration) and perform informed search, (2) a random dispersal behavior where individuals randomly leave and randomly search for a new colony and (3) a semi-informed dispersal behavior where individuals leave only poor habitat colony (informed emigration) but settle randomly among reachable colonies (randomly search).

Migration matrix M. For the **random emigration behavior**, the emigration rate m_i for each colony i depends only on the proportion p_{mi} of individuals that leave the colony. Thus the migration matrix is a diagonal matrix with diagonal coefficient $m_i = p_{mi}$.

For the **informed emigration behavior**, the emigration rate depends on the quality of the habitat, measured through the realized population growth $r_i^*(t)$ and a sensitivity parameter $\mathbf{p}_m = (p_{m1}, \ldots, p_{m7})$ measuring the intensity of the emigration. The realized population growth rate $r_i^*(t)$ is a function of both the intrinsic growth rate $r_i(t)$, and the carrying capacity of the colonies K_i:

$$r_i^*(t) = \begin{cases} (1 + r_i(t)) \exp\left(1 - \dfrac{N_i(t)}{K_i}\right) - 1 & \text{if } r_i(t) > 0 \\ r_i(t) & \text{if } r_i(t) \leqslant 0 \end{cases}$$

We assume that the emigration rate m increases linearly from $m = 0$ at $r^* \geqslant 0$ to $m = 1$ at critical value $r_c^* < 0$. Thus, a critical threshold r_c^* close to 0 corresponds to high migration, while a larger threshold reflects low migration. Here, we will estimate the emigration rate parameter $\mathbf{p}_m = (p_{m1}, \ldots, p_{m7})$, which quantifies the critical value in each region i:

$$r_{i,c}^* = (1 - p_{mi}) r_m^*$$

where r_m^* is the lowest intrinsic growth rate. Hence, the emigration matrix $\mathbf{M}$ only depends

on the ratio $\mathbf{r}^*(t)/(1-\mathbf{p}_m)r_m^*$

$$m_i\left[\frac{r_i^*(t)}{(1-p_{mi})r_m^*}\right]=\begin{cases}1 & \text{if } r_i^*(t)<(1-p_{mi})r_m^*\\ 1-\dfrac{r_i^*(t)}{(1-p_{mi})r_m^*} & \text{if } (1-p_{mi})r_m^*\leqslant r_i^*(t)\leqslant 0\\ 0 & \text{if } r_i^*(t)>0\end{cases} \tag{3}$$

Searching matrix S. Once individuals have left their colonies, we assume that they search for a new colony. The searching matrices $\mathbf{S}[\mathbf{x}]$ is

$$\mathbf{S}_{ij}[\,\mathbf{x},d\,]:=\mathcal{S}(j|i,\mathbf{x}(t),d),\quad \text{for } j\neq i \quad\text{and}\quad \mathbf{S}_{ii}[\,\mathbf{x},d\,]:=-\sum_{j\neq i}\mathbf{S}_{ij}[\,\mathbf{x},d\,],$$

indicating that the probability of settlement in a colony j depends on leaving colony i, the characteristics of the habitat in the colony j ($\mathbf{x}(t)$), and the dispersal ability of the individuals d.

For the random search individuals can move randomly across landscape according to a dispersal kernel $k(x)$ which describes the probability of traveling a distance x. This probability distribution can take various forms according to the dispersal ability of the species. In our simulation, we use a uniform kernel because emperor penguins have the ability to cover incredible distances, thus all colonies are potentially connected:

$$k_{unif}(x):=\frac{1}{d}\mathbf{1}_{[0,d]}(x),\quad \text{for all } x\in[0,+\infty),$$

where d represents the mean distance dispersal of the species and $\mathbf{1}_{[0,d]}(x)$ is the characteristic function of the interval $[0,d]$. Thus under the random search, the probability $S_R(j|i,\mathbf{x}(t),d)$

of moving to colony j given that individual left its resident colony i at time t is defined by

$$S_R(j|i, \mathbf{x}(t), d) := \frac{k(\text{dist}(i,j))}{\sum_{j \neq i} k(\text{dist}(i,j))}. \tag{4}$$

where $\text{dist}(i,j)$ corresponds to the landscape topography, specifically the coastal distance between colonies in our case study. With the random search, individuals may settle in a new colony of lower quality than their resident colony.

A.3 Computation of Likelihood

Conditional probability of genotype $\mathcal{G}_{i,\tau}$ The computation of the genotype likelihoods involves classical genetic assignment approaches Paetkau *et al.* [1995]; Pritchard *et al.* [2000] and dispersal analysis from spatially sampled data. In each sampled colonies τ we have captured and genotyped G_τ individuals. The number of genotyped individuals G_τ ranges from 10 to 16 individuals depending on the sampled colony τ. Thus we have G_τ genotypes $\mathcal{G}_{i,\tau}$ with $i = 1, \ldots, G_\tau$.

The conditional probability that an individual i carries alleles $(a^1, a^2) \in \{1, \ldots, A_\lambda\}^2$ at locus λ given that this individual comes from cluster r, can be deduced from the allele frequencies of each cluster $\mathcal{F}_{r\lambda} = (p_{r\lambda a})_{a=\{1,4\}}$. From the Hardy–Weinberg equilibrium assumption within a cluster, the two-alleles are independent and alleles frequencies are:

$$\mathbb{P}\big((a^1, a^2)|\text{ indiv. } i \text{ belongs to cluster } r\big) = 2^{k_\lambda} p_{r\lambda a^1} p_{r\lambda a^2} \tag{5}$$

where $k_\lambda = 0$ if the individual is homozygous at locus λ, that is $a^1 = a^2$, and $k_\lambda = 1$ otherwise. Using the linkage equilibrium among loci, the conditional probability for the

genotype $\mathcal{G}_{i\tau}$ is:

$$\mathbb{P}\big(\mathcal{G}_{i,\tau}|\text{ indiv. } i \text{ belongs to cluster } r\big) = 2^{k_i}\prod_{\lambda=1}^{\Lambda} p_{r\lambda a^1}p_{r\lambda a^2} \tag{6}$$

where k_i is the number of heterozygous loci in $\mathcal{G}_{i,\tau}$.

Probability of sampling individual from a cluster Since individuals may move from colony to colony, we need to model the genetic sampling of individuals in the different colonies. The sampling of individuals at a given time t at a colony τ is random among the individuals observed at the colony. Thus the expected number of individuals that can be potentially captured $C_\tau(t)$ is proportional to the number of individuals alive in the colony τ at time t:

$$C_\tau(t) = \beta_\tau n_\tau(t),$$

where β_τ is the capture rate at colony τ and $n_\tau(t)$ follows the dynamics described by equation (1).

However, within a colony, the population is structured into the different neutral genetic clusters. The number of individuals within colony i from cluster r is denoted by $n_i^r(t)$ and the total number of individuals, which belong to the cluster r across colonies is $\mathbf{n}^r(t)$. Since we are looking at neutral set of loci, all individuals are supposed to share the same dispersal and reproduction characteristics independently of their clusters. Thus the number of individuals $\mathbf{n}^r(t)$ satisfies [Roques *et al.*, 2012]:

$$\mathbf{n}^r(t+1) = \mathbf{D}\big[t, \mathbf{n}(t)\big]\mathbf{F}\big[t, \mathbf{n}(t)\big]\,\mathbf{n}^r(t) \tag{7}$$

and initially,

$$n_i^r(0) = \mu_i^r n_i(0), \quad \text{for all } i \in \{1, \ldots, 54\}, \tag{8}$$

where μ_i^r is the initial proportion of the cluster r within a colony i ($\sum_{r=1}^{R} \mu_i^r = 1$ for all i). The genetic and demographic dynamics are linked by:

$$\mathbf{n}(t) = \sum_{r=1}^{R} \mathbf{n}^r(t). \tag{9}$$

Thus, the expected number of individuals $C_\tau^r(t)$ belonging to a cluster r that can be sampled at colony τ at time t is given by

$$C_\tau^r(t) = \beta_\tau n_\tau^r(t). \tag{10}$$

Since the genotyping process corresponds to a sampling without replacement, the number of genotyped individuals in τ belonging to cluster r follows a multivariate hypergeometric distribution with parameters C_τ, C_τ^r and G_τ, the sample size. In our data collection, the number G_τ of sampled individuals per colony ranges from 10 to 16. From previous studies Jenouvrier *et al.* [2020], the estimated number of individuals in the colonies, where data have been collected, ranges from 213 to 22510 in 2009. Thus C_τ is quite large and it is larger than the average number of sampled individuals per colony. Thus the multivariate hypergeometric distribution can be approximated by a multinomial distribution with parameters G_τ and $(C_\tau^1/C_\tau, \ldots, C_\tau^R/C_\tau)$. Using this approximation, we recover the probability that a genotyped individual i observed in colony τ belongs to cluster r, stated in the main text:

$$\mathbb{P}\big(\text{ indiv. } i \text{ belongs to } r\big) = \frac{C_\tau^r}{C_\tau} = \frac{n_\tau^r(t_i)}{n_\tau(t_i)} = \mu_\tau^r(t_i). \tag{11}$$

where $\mu_\tau^r(t_i)$ is the frequency of the genetic cluster r in the colony τ at time t_i.

Likelihood function We deduce the probability of the genotype $\mathcal{G}_{i,\tau}$

$$\mathbb{P}(\mathcal{G}_{i,\tau}) = \sum_{r=1}^{R} \mu_\tau^r(t_i) \left[2^{k_i} \prod_{\lambda=1}^{\Lambda} p_{r\lambda a^1} p_{r\lambda a^2} \right] \tag{12}$$

Finally, we compute the likelihood function associated with the unknown parameters, given the genotyped $\mathcal{G}_{i,\tau}$ and the sampling times (t_i) as:

$$\begin{aligned}\mathcal{L}(\Theta) &= \prod_{\tau=1}^{J}\prod_{i=1}^{G_\tau}\mathbb{P}(\mathcal{G}_{i,\tau}|\Theta) \\ &= \prod_{\tau=1}^{J}\prod_{i=1}^{G_\tau}\sum_{r=1}^{R}\mu_\tau^r(t_i)\left[2^{k_i}\prod_{\lambda=1}^{\Lambda}p_{r\lambda a^1}p_{r\lambda a^2}\right]\end{aligned} \tag{13}$$

A.4 Estimation of parameters

We aim to infer the following parameter vector $\Theta = (d, \mathbf{p}_m, \boldsymbol{\mu})$.

First, we assume that the vector of parameters $\mathbf{p}_m$ has only 7 different values corresponding to the seven geographical regions defined in section A.1. In addition, the initial proportions $\boldsymbol{\mu}$ of clusters are only unknown in the colonies belonging to the three regions without genetic characterisation (StoS, StoK and A-B sea).

For each behavior (informed, random or semi-informed see section A.2), the estimates of dispersal parameters $(\widehat{d}, \widehat{\mathbf{p}}_m)$ and the initial cluster proportions $(\widehat{\boldsymbol{\mu}})$ have been obtained by minimizing the logarithm of the inverse Likelihood, that is $-\log(\mathcal{L}(d, \mathbf{p}_m, \boldsymbol{\mu}))$.

The minimization algorithm is performed using a Bayesian method, where the prior of the parameters d and $\mathbf{p}_m$ are assumed to be uniformly distributed with the following constraints:

$$(\widehat{d}, \widehat{\mathbf{p}}_m) \in (250, 6500) \times (0,1)^7$$

and the prior of the parameter $\boldsymbol{\mu}$ follows a Dirichlet distribution of order $R = 4$ with parameters all equal to 1, thus we have:

$$\widehat{\mu}_h^r \in (0,1)^4 \quad \text{and} \quad \sum_{r=1}^{4}\widehat{\mu}_h^r = 1 \quad \text{for all } h \in \{1, \ldots, 54\}.$$

A.5 Model selection criteria

We explore 3 dispersal behaviors (informed, random or semi-informed see section A.2). In order to choose the most likely behavior, we perform a model selection using the different selection model criteria: the Bayesian Information Criteria (BIC) Schwarz [1978], two Deviance Information Criteria (DIC) Gelman *et al.* [2003]; Spiegelhalter *et al.* [2002] and a predictive Information Criteria (IC) Ando [2011]. The BIC is defined by

$$BIC = -2\log[\mathcal{L}(\Theta^*)] + k\log(I) \tag{14}$$

where I is the sample size, k the number of parameters and Θ^* is the maximum likelihood estimate of the parameter vector Θ, that is $\Theta^* = \text{argmax}(\mathcal{L}(\Theta))$. In our study, k and I are the same for all the models.

The DIC satisfies

$$DIC = \widehat{\mathcal{D}} + p_{eff} \tag{15}$$

where $\widehat{\mathcal{D}}$ is the posterior mean of the deviance $\mathcal{D}(\Theta) = -2\log[\mathcal{L}(\Theta^*)]$ and p_{eff} is the effective number of parameters of the model. We use two different versions of the DIC, which rely on different definitions of p_{eff}. The first version has been developed by Spiegelhalter *et al.* [2002]:

$$p_{eff} = \widehat{\mathcal{D}} - \mathcal{D}(\widehat{\Theta}) \tag{16}$$

where $\widehat{\Theta}$ is the posterior mean of Θ. The second version has been introduced by Gelman *et al.* [2003]:

$$p_{eff} = \frac{1}{2}\mathbb{V}(\mathcal{D}(\Theta)) \tag{17}$$

where $\mathbb{V}(\mathcal{D}(\Theta))$ is the posterior variance of $\mathcal{D}(\Theta)$.

We also use the IC developed by Ando [2011]:

$$IC = 3\widehat{\mathcal{D}} - 2\mathcal{D}(\widehat{\Theta}) \tag{18}$$

In practice, the posterior mean and variance, which appear in our four criteria, are computed with their empirical values using the weighted posterior sample $\{\Theta_m, w_m\}$ provided by our minimization algorithm.

B Estimated parameters

B.1 Estimation of the initial proportion of cluster within a colony.

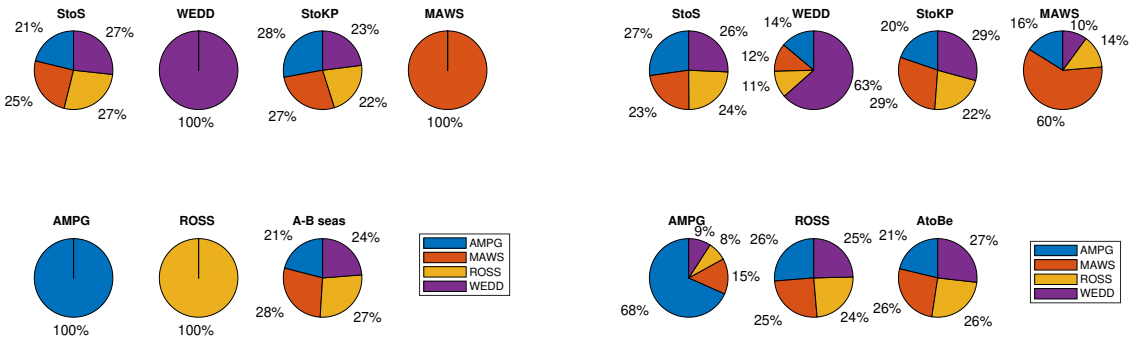
The model permits to estimate the initial proportion of cluster within colonies. Since the clusters are associated to geographical regions (AMPG, WEDD, ROSS and MAWS), we assume that for the colonies belonging to those regions, the initial proportion of the corresponding cluster is equal to 1 (Fixed observed cluster on Fig. S1).

For the semi-informed dispersal (Fig. S1), we found that the colonies

- from Snowhill to Smith (StoS) are mostly composed of individuals from WEDD cluster (geographically closest cluster) and AMPG cluster.
- from Stancomb to Kloa Point (StoKP) are mostly composed of MAWS and WEDD clusters;
- from Ledda bay to Rotschild (AtoBe) seems panmictic because all the proportion are almost equal.

This pattern is confirmed even when we do not constrain the initial proportion of the observed cluster (Estimated observed cluster on Fig. S1b). In that case the highest proportion of a

given cluster matches with the observed cluster. This correct assignment suggests that our model reproduces well the expected pattern.

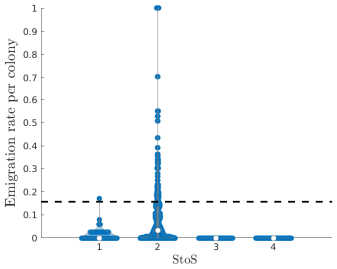


(a) Fixed observed cluster – Semi-Informed behavior (b) Estimated observed cluster – Semi-Informed behavior

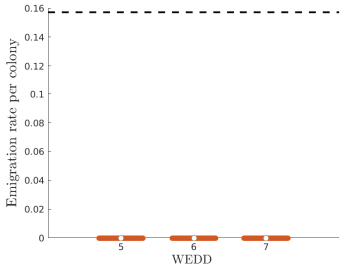
Figure S1: Estimation of the proportion of the four genetic clusters (AMPG, MAWS, ROSS and WEDD) across the seven geographical regions in Antarctica.

B.2 Emigration rate per colony

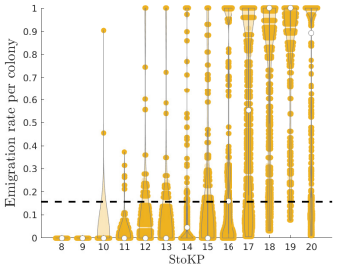
In the main text we gather colonies per geographical regions. However, we quantify the emigration rates for each colony (see Fig. S2). We observe that the emigration rates vary among colonies belonging to a similar geographical regions. In addition, we provide the averaged migration flux between colonies of Antarctica from 2009 to 2014 (Fig. S3) for the following regions: Smith to Snowhill Island in the Wedell sea (StoS), Weddell sea (Gould Bay to Halley Bay colonies) (WEDD), from Stancomb to Kloa point (StoK), Mawson Bay (Fold Island to Cape Darnley colonies) (MAWS), from Amanda Bay to Pointe Geologie colonies (AMPG), the Ross sea (Cape Washington and Cape Crozier) (ROSS) and Admunsen and Bellinghausen seas (Ledda bay to Rothschild Island) (A-B seas)



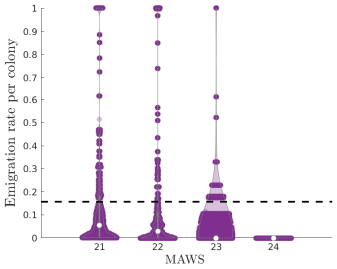
(a) Smith to Snowhill Island (StoS)



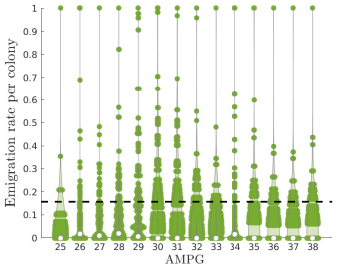
(b) Weddell sea (WEDD)



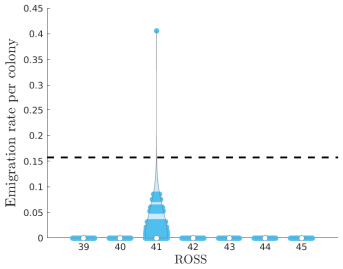
(c) Stancomb to Kloa Point colonies (StoKP)



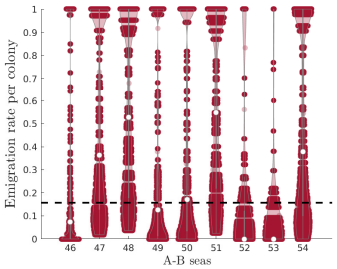
(d) Mawnson bay (MAWS)



(e) Amanda Bay to Pointe Geologie (AMPG)



(f) Ross sea (ROSS)



(g) Admunsen and Bellinghausen seas (A-B seas)

Figure S2: Posterior distributions of the emigration rates per colony over the entire Antarctic continent (a) and for the seven regions of Antarctica (b). White dots represent the median of the distributions and the black line is the mean emigration rate over all colonies in Antarctica (0.157)

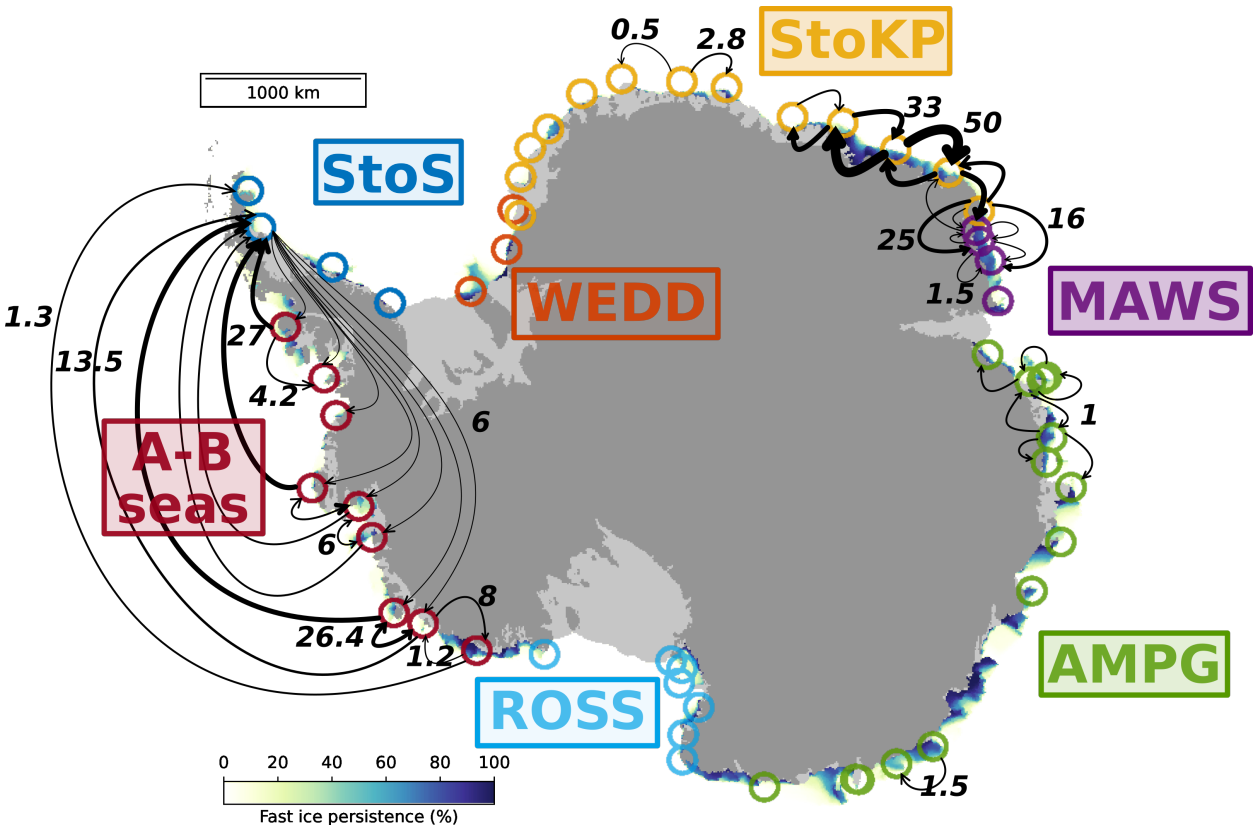


Figure S3: Migration flux estimation averaged between colonies of Antarctica from 2009 to 2014. Map background shows the distribution of fast ice persistence, expressed as a percentage of time covering each unit (square of 6.25km) from March 2000 to March 2018 (the map is extracted from Fraser *et al.* [2021]).

C Importance of climatic and demographic covariates

C.1 Description of the demographic and environmental factors

We use three demographic factors independent from our meta-population analysis: the size of colony, the growth rate per colony and the frequency of blinking that corresponds to the relative number of year a colony disappear over a period of 10 years, from 2009 to 2018. They were calculated from unpublished data of colony presence and population counts of emperor penguins from VHR satellite imagery [LaRue *et al.*, 2022]. Those estimates represent the portion of the colony in attendance on the fast ice every year and thus available for surveying through aerial counts or satellite images during the chick-rearing season [LaRue *et al.*, 2022].

We also considered different environmental variables around each colony: the zooplankton biomass (mmol C/m^2) and the distance between the colony and the nearest edge of fast ice (m), the fast ice area around a colony (m^2), the emergence and breaking date of the fast ice, the mixed layer depth (m), the upper ocean temperature (top 10m,$^{\circ}C$), and surface wind (m/s). We also used new fast ice variables from a recent analysis of Emperor penguin habitat Labrousse *et al.* [2021] describing the persistence and the magnitude of fast ice annual cycle. Fast ice variables were computed using continuous, high-spatio-temporal resolution time series of circum-Antarctic fast-ice extent from [Fraser *et al.*, 2020]. Other reanalysis environmental products were computed in a forced ocean sea ice (FOSI) configuration of the Community Earth System Model (CESM2, 1^o resolution) [Danabasoglu *et al.*, 2020]. Additionally, we used different scales: the average within a 100 km buffer around each colony when calculating fast ice variables; 800 km buffer for other variables during the non-breeding period; and a 500 km buffer for other periods.

C.2 Importance variable analysis through random forest algorithm

First, we modeled the proportion of emigration years across colonies with both demographic variables and environmental variables using conditional random forests [Strobl *et al.*, 2007]. The best models had an R^2 value of 0.48.

Second, we refine our analysis using environmental variables only by modeling the annual emigration probability using the same framework of conditional random forests. The best model had an area under the receiving operator characteristic curve (AUC) of 0.82, which was calculated with a 10-fold cross validation, signifying good classification performance. Variable importance scores were calculated with conditional permutation importance using the R package "permimp" [Debeer & Strobl, 2020] (see Figure S4). The codes are available online https://github.com/bilgecansen/Emperor_dispersal.

Here we present the figure of the conditional variable importance scores of random forests modelling annual emigration probability when refining our analysis only with environmental covariates (Fig. S4).

C.3 Distribution analysis

In addition, we show the distribution of the size of the colonies for colonies with or without emigration (Fig. S5).

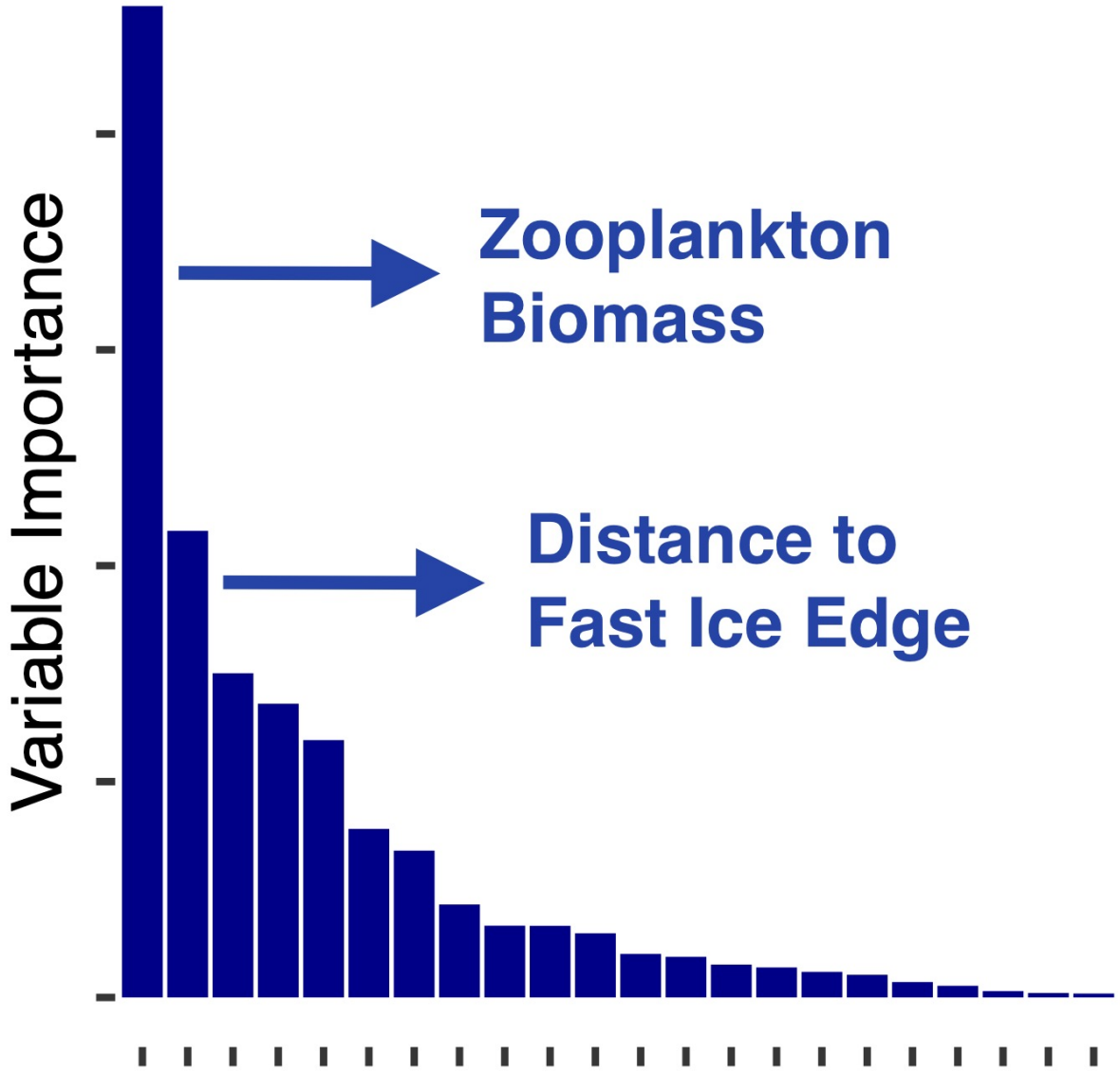


Figure S4: Conditional variable importance scores of random forests modelling annual emigration probability. Only the top 2 variables are shown.

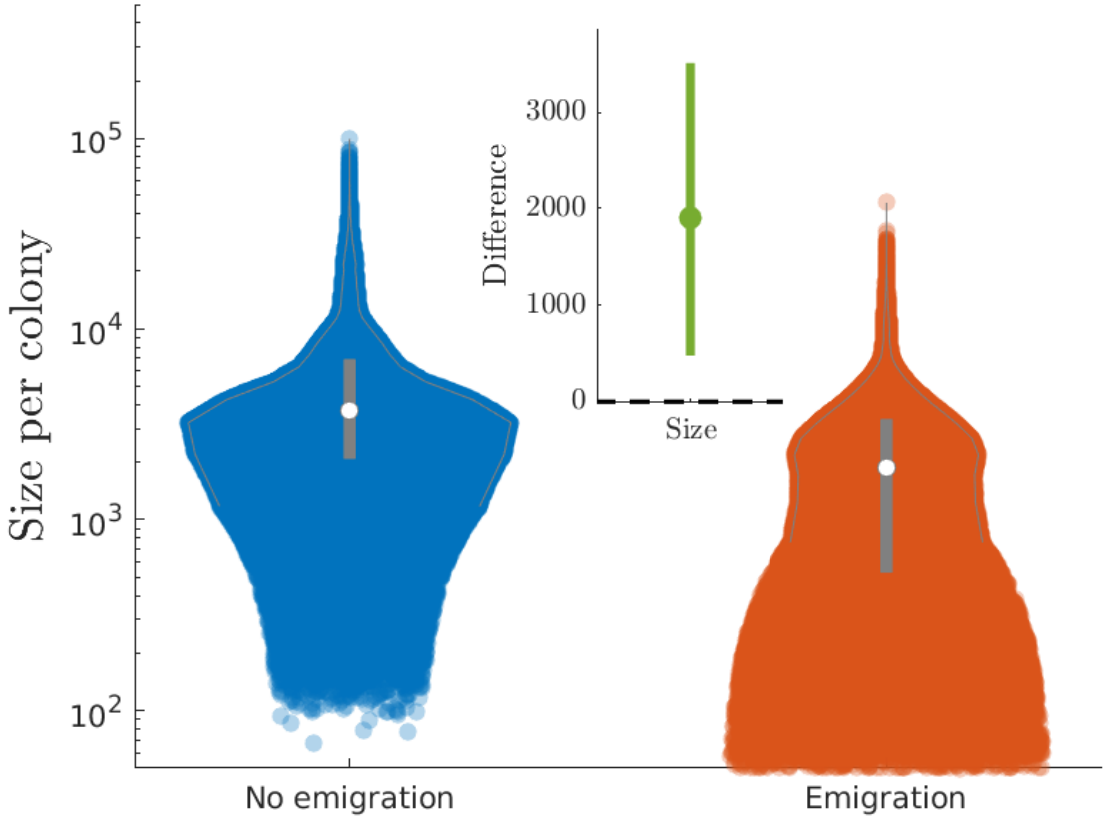
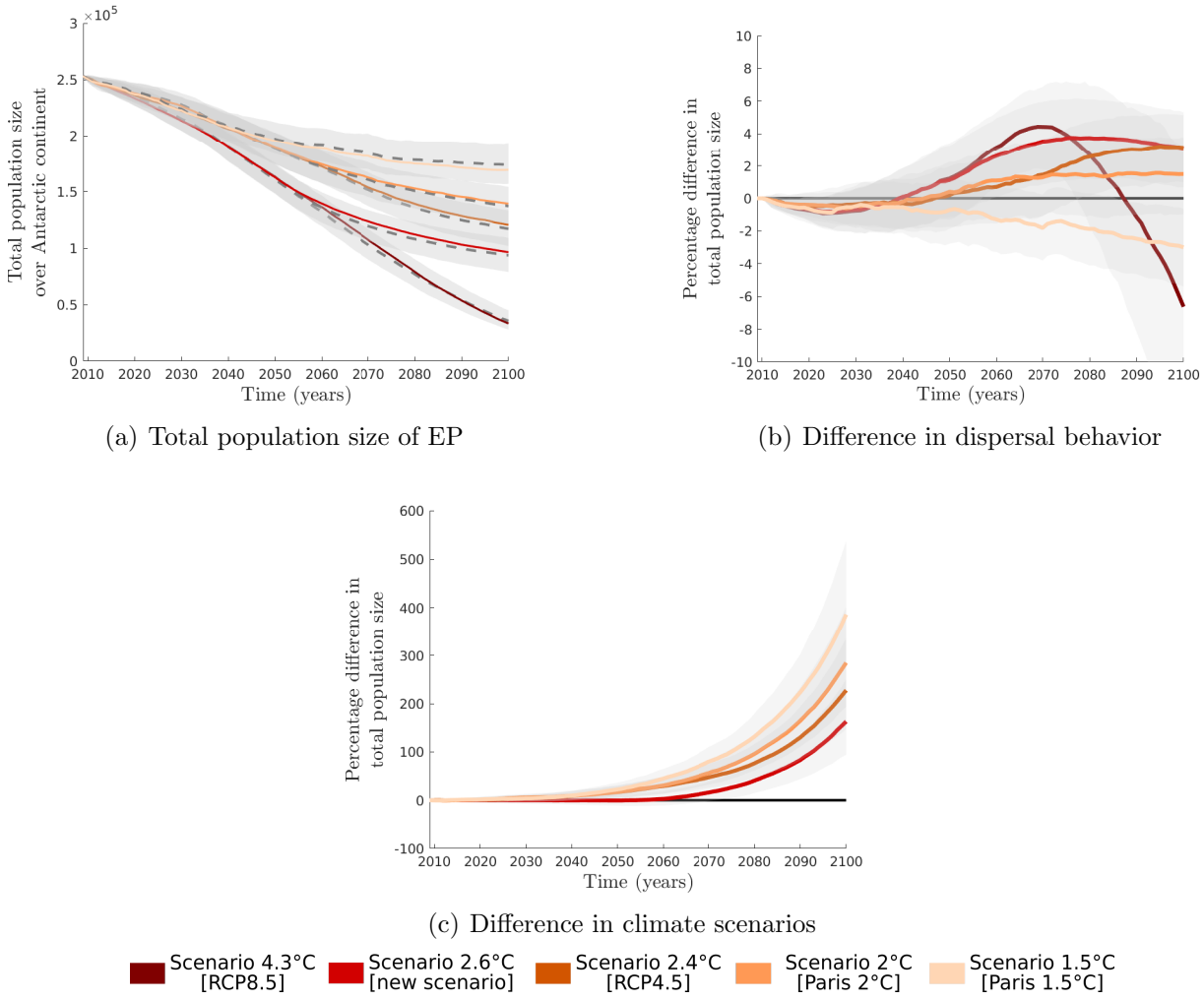


Figure S5: Distribution of the size per colony for colonies with Emigration (orange violinplots) and without emigration (blue violinplots). The subparts represent the boxplot of the difference between demographic covariates in colonies without and with emigration (green corresponds to positive difference). White dots correspond to the median of the distributions.

D Projection of the total Emperor penguin population size

Plugging the estimated demographic parameters provided by our new analysis into our metapopulation model (section A.2), we project the total population size of Emperor penguin over the century in different climate scenarios: scenario 4.3°C [RCP8.5], scenario 2.6°C [new scenario], scenario 2.4°C [RCP4.5], scenario 2°C [Paris 2°C] and scenario 1.5°C [Paris 1.5°C]. We compare the outcome of this updated model (semi-informed dispersal), with the projections of the model without dispersal (see Fig. S6). We show that the more likely dispersal dispersal behavior predicted for the Emperor penguins (semi-informed dispersal with small mean distance dispersal 428km and small emigration rates), results in a greater global population up to 5% compared with a scenario without dispersion when the climate scenarios are unfavorable (from scenario 4.3°C and scenario 2°C). However, under a favorable climate scenario 1.5°C (Paris 1.5°C), this dispersal behavior does not improve significantly the global population size but may attenuate it. We also compared our projection with the inferred dispersal behavior with the projections obtain with a broader range of dispersal behaviors (random, semi-informed or informed dispersal) and various mean dispersal distances and emigration rates. We show that our predicted scenario is not among the more optimistic for the Emperor penguin population size because we project on average a 5% reduction of the population size compared with the other scenario of dispersal. As a conclusion, the impact of dispersal behavior, distance and emigration rate on the future global population size is relatively small compared with the impact of climate change mitigation Jenouvrier *et al.* [2021, 2020] (see Fig.SI6 and SI7).



(a) Total population size of EP

(b) Difference in dispersal behavior

(c) Difference in climate scenarios

Figure S6: Projection of the total population size of Emperor penguin from 2009 to 2100 using the demographic model (1) with different climate scenarios. In panel (a), the projected total population size without dispersal (dashed curves) and with semi-informed dispersal with the most likely parameters provided by our analysis (plain curves); The grey regions corresponds to the 1% confident intervals around the median. In panel (b), we present the percentage difference of population size between the projection with a semi-informed dispersal and the projection without dispersal for each climate scenario. In panel (c), we present the percentage difference of population size between the projection with the worst climate scenario 4.3°C [RCP8.5] and the other scenarios.

References

Ando T (2011) Predictive bayesian model selection. *American Journal of Mathematical and Management Sciences*, **31**, 13 – 38.

Danabasoglu G, Lamarque J, Bacmeister J, *et al.* (2020) The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, **12**.

Debeer D, Strobl C (2020) Conditional permutation importance revisited. *BMC Bioinformatics*, **21**. doi:10.1186/s12859-020-03622-2.

Fraser AD, Massom RA, Handcock MS, *et al.* (2021) Eighteen-year record of circum-antarctic landfast-sea-ice distribution allows detailed baseline characterisation and reveals trends and variability. *The Cryosphere*, **15**, 5061–5077. doi:10.5194/tc-15-5061-2021.

Fraser AD, Massom RA, Ohshima KI, Willmes S, Kappes PJ, Cartwright J, Porter-Smith R (2020) High-resolution mapping of circum-antarctic landfast sea ice distribution, 2000–2018. *Earth System Science Data*, **12**, 2987–2999.

Gelman A, Carlin J, Stern H, Dunson D, Vehtari A, Rubin D (2003) *Bayesian Data Analysis*. Chapman & Hall / CRC.

Jenouvrier S, Che-Castaldo J, Wolf S, *et al.* (2021) The call of the emperor penguin: Legal responses to species threatened by climate change. *Global change biology*, **27**, 5008–5029.

Jenouvrier S, Garnier J, Patout F, Desvillettes L (2017) Influence of dispersal processes on the global dynamics of emperor penguin, a species threatened by climate change. *Biol. Conserv.*, **212**, 63 – 73. doi:10.1016/j.biocon.2017.05.017.

Jenouvrier S, Holland M, Iles D, *et al.* (2020) The paris agreement objectives will likely halt future declines of emperor penguins. *Global change biology*, **26**, 1170–1184.

Labrousse S, Fraser AD, Sumner M, *et al.* (2021) Landfast ice: a major driver of reproductive success in a polar seabird. *Biology Letters*, **17**, 20210097.

LaRue M, Brooks C, Wege M, Salas L, Gardiner N (2022) High-resolution satellite imagery meets the challenge of monitoring remote marine protected areas in the antarctic and beyond. *Conservation Letters*, **15**, e12884. doi:https://doi.org/10.1111/conl.12884.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in canadian polar bears. *Mol. Ecol.*, **4**, 347–354. doi:10.1111/j.1365-294X.1995.tb00227.x.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959. doi:10.1093/genetics/155.2.945.

Roques L, Garnier J, Hamel F, Klein EK (2012) Allee effect promotes diversity in traveling waves of colonization. *Proc Natl Acad Sci USA*, **109**, 8828–8833.

Schwarz G (1978) Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, 583–639. doi:https://doi.org/10.1111/1467-9868.00353.

Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, **8**. doi:10.1186/1471-2105-8-25.

Younger JL, Clucas GV, Kao D, Rogers AD, Gharbi K, Hart T, Miller KJ (2017) The challenges of detecting subtle population structure and its importance for the conservation of emperor penguins. *Molecular Ecology*, **26**, 3883–3897. doi:10.1111/mec.14172.