

# Analytical Mixing Model, v2

Mari Kawakatsu

Last updated: February 12, 2019

## Contents

<b>1 Overview</b>	<b>1</b>
<b>2 Model</b>	<b>1</b>
<b>3 Steady-state predictions</b>	<b>2</b>
3.1 Pure colonies: 2 tasks, 1 line . . . . .	2
3.2 Mixed colonies: 2 tasks, 2 lines, 50-50 mixes . . . . .	3
3.2.1 Comparing with simulation results . . . . .	4
3.3 Mixed colonies: 2 tasks, 2 lines, non-50-50 mixes (NEW) . . . . .	5
3.3.1 Comparing with simulation results . . . . .	5
<b>4 Downward vs. upward contagion in the 50-50 mixes</b>	<b>6</b>
<b>5 Questions</b>	<b>7</b>

## 1 Overview

- Following up on the last Skype with Yuko & Daniel, I ran additional simulations in which I varied a combination of parameters, namely the task efficiencies  $\alpha$  and the demand rates  $\delta$  (results on slides).
- There seemed to be some strange behavior when at least one of the lines was inefficient. To understand what was happening, I decided to write down equations that approximate the computational model to gain insight into the (expected) steady-state behavior (see Sections 2 and 3).
- A preliminary analysis of the model seems to suggest that, assuming that the system reaches/has a steady state, varying only  $\delta$  and  $\alpha$  can result in a downward contagion but not an upward contagion (see Section 4). If this is true, then in order for the FTM to capture the experimental data, it must be the case that the two genetic lines differ in some trait other than their task efficiency.
- (Update 2/9) I extended the model to capture non-50-50 mixes.

## 2 Model

To consider the dynamics of division of labor in mixed colonies, we extend the fixed threshold model in [1] to incorporate ants of two genetic lines, **A** and **B**.

The model considers  $n$  individuals and  $m$  tasks. Let  $f$  and  $1 - f$  be the fractions of genetic line **A** and **B** individuals in the colony, respectively. We assume that each task  $j$  has an associated stimulus,  $s_{j,t}$ , at every time  $t$ , indicating group-level demand for that task. We model the change in stimulus over discrete time as

$$s_{j,t+1} - s_{j,t} = \delta_j - \frac{\alpha_j^A x_{j,t}^A + \alpha_j^B x_{j,t}^B}{n}, \quad (1)$$

where

- $\delta_j$  is the task-specific demand rate;

- $\alpha_j^A$  and  $\alpha_j^B$  are the task-specific performance efficiencies of lines A and B, respectively; and
- $x_{j,t}^A$  and  $x_{j,t}^B$  are the numbers of line A and line B individuals performing task  $j$  and time  $t$ , respectively.

At each time step, inactive individuals are exposed to the task stimuli randomly until either they begin performing a task or they have encountered all stimuli without landing on a task. Similar to [1], our computational model draws individual  $i$ 's internal threshold for task  $j$ ,  $\theta_{ij}$ , from a normal distribution with mean  $\mu_j$  and normalized standard deviation  $\sigma_j$  (each of these parameters can be line- and/or task-specific). *To gain analytical insight into the model, we make the simplifying assumption that the line- and task-specific thresholds are given by the constant parameters,  $\mu_j^A$  and  $\mu_j^B$ .*

With this assumption, the probabilities  $P_{ij,t}^A$  and  $P_{ij,t}^B$  that individuals  $i$  of lines A and B, respectively, perform task  $j$  at time  $t$  can be written as

$$P_{j,t}^A(s_{j,t}) = \frac{s_{j,t}^\eta}{s_{j,t}^\eta + (\mu_j^A)^\eta}, \quad P_{j,t}^B(s_{j,t}) = \frac{s_{j,t}^\eta}{s_{j,t}^\eta + (\mu_j^B)^\eta}, \quad (2)$$

where  $\eta$  is the threshold response parameter as in the original model. Finally, at a given time  $t$ , active individuals quit their tasks with a constant quit probability  $\tau$ . In the case of two tasks ( $m = 2$ ), the change in the number of line A and line B individuals working on task  $j$  is given by<sup>1</sup>

$$\begin{aligned} x_{j,t+1}^A - x_{j,t}^A &= \frac{1}{2} \left[ P_{j,t}^A(s_{j,t}) + (1 - P_{j',t}^A(s_{j',t}))P_{j,t}^A(s_{j,t}) \right] \left( fn - (x_{j,t}^A + x_{j',t}^A) \right) - \tau x_{j,t}^A \\ x_{j,t+1}^B - x_{j,t}^B &= \frac{1}{2} \left[ P_{j,t}^B(s_{j,t}) + (1 - P_{j',t}^B(s_{j',t}))P_{j,t}^B(s_{j,t}) \right] \left( (1-f)n - (x_{j,t}^B + x_{j',t}^B) \right) - \tau x_{j,t}^B \end{aligned} \quad (3)$$

for  $(j, j') = (1, 2), (2, 1)$ . The sums in square brackets capture the possible ways in which individuals can initiate task  $j$ : they can either encounter the stimulus for task  $j$  immediately and begin performing that task, or they can first encounter the stimulus for the other task  $j'$ , decide not to perform that task, subsequently encounter the stimulus for task  $j$ , and begin performing task  $j$ .

The full model for  $m = 2$  consists of the six equations describing changes in  $s_1, s_2$  (Eq. (1)) and  $x_1^A, x_1^B, x_2^A, x_2^B$  (Eq. (3)), under the constraints  $x_{1,t}^A + x_{2,t}^A \leq fn$  and  $x_{1,t}^B + x_{2,t}^B \leq (1-f)n$ .

### 3 Steady-state predictions

One way to understand what is happening in the simulations is to consider the equilibria of the above model. My intent here is not (yet) to prove the stability etc. rigorously but to get a sense for what we should expect to see in the simulations.

#### 3.1 Pure colonies: 2 tasks, 1 line

For pure (A-only) colonies, we set  $x_{j,t}^B = 0$  for all time  $t$ . Setting Eq. (1) to zero, we obtain

$$\frac{x_j^A}{n} = \frac{\delta_j}{\alpha_j^A} \quad (4)$$

as the fraction of (A) ants that are performing task  $j$  ( $j = 1, 2$ ) at steady state. We can use this to compute the steady-state levels of stimuli if needed.

**Notably, the steady-state values of  $x_j^A$  are independent of the mean threshold ( $\mu_j^A$ ) or the quit probability ( $\tau^A$ ).** This agrees with the results of our simulations, where varying  $\mu$  and  $\tau$  by line did not lead to differing mean task performance levels in the pure colonies.

Note that the sum of the fractions (4) across tasks should not exceed one. So a necessary condition for the steady-state to be biologically realizable is

$$\frac{x_1^A}{n} + \frac{x_2^A}{n} = \frac{\delta_1}{\alpha_1^A} + \frac{\delta_2}{\alpha_2^A} \leq 1. \quad (5)$$

---

<sup>1</sup>explain this more in English

If this condition is not met, then we would expect the stimuli to continue growing (i.e., the system will not reach a steady state). A “symptom” of this phenomenon, I think, is that the ants become “maxed out” (I can explain this more when we Skype).

### 3.2 Mixed colonies: 2 tasks, 2 lines, 50-50 mixes

We first consider the 50-50 mixes ( $f = 0.5$ ), i.e., mixed colonies consisting of an equal number of **A** and **B** ants. Moreover, we assume that the mean thresholds and the quit probabilities are identical for both tasks and lines ( $\mu_1^A = \mu_2^A = \mu_1^B = \mu_2^B$  and  $\tau^A = \tau^B$ ), as we have done in the simulations with varied  $\delta$  and  $\alpha$  values. Setting Eq. (1) and Eq. (3) equal to zero, we find<sup>2</sup> the expression for the steady-state numbers of individuals performing task  $j$ :

$$x_j^A = x_j^B = n \left( \frac{\delta_j}{\alpha_j^A + \alpha_j^B} \right) \quad (6)$$

for  $j = 1, 2$ . Alternatively, as fractions of each genetic type of individuals,

$$\frac{x_j^A}{(n/2)} = \frac{x_j^B}{(n/2)} = \frac{2\delta_j}{\alpha_j^A + \alpha_j^B}. \quad (7)$$

The necessary condition for this steady state to exist is

$$\sum_{j=1}^2 \frac{x_j^A}{n} + \frac{x_j^B}{n} = \sum_{j=1}^2 \frac{2\delta_j}{\alpha_j^A + \alpha_j^B} \leq 1. \quad (8)$$

Again, if this condition is not met, then we would expect the stimuli to continue growing over time and for the ants to be working at max capacity.

While  $\mu$  doesn't explicitly appear in (7) when we assume that the mean thresholds are identical for all individuals and both tasks, we expect<sup>3</sup> the general form of steady state fractions of active individuals to be functions of  $\mu_j^A$  and  $\mu_j^B$  as well as  $\tau^A$  and  $\tau^B$ . **So one question to investigate using simulations (numerical or computational) would be how the mean thresholds ( $\mu$ ) interact with either the demand rates ( $\delta$ ) or the task efficiencies ( $\alpha$ ).**

<sup>2</sup>This is the expression I get when I set  $1 \leq \eta \leq 20, \eta \in \mathbb{Z}$ . There seems to be something numerically tricky when  $\eta$  takes non-integer values.

<sup>3</sup>I haven't been able to get Mathematica to compute the expressions explicitly, but we can infer this from the form of the threshold functions.

### 3.2.1 Comparing with simulation results

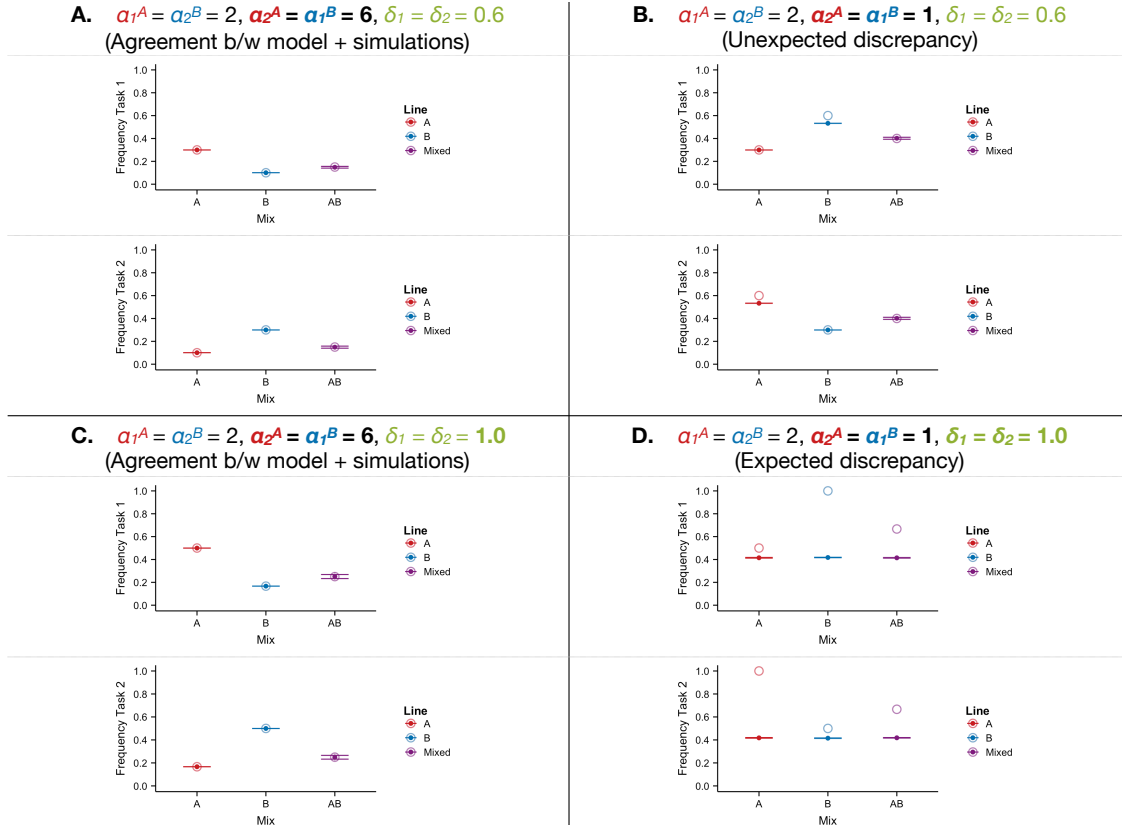


Figure 1: Comparison of simulation results (solid dots with error bars) and analytical predictions of steady states (opaque rings). The following parameters are shared among the four cases (A-D):  $\mu = 10, \sigma = 0.1, \eta = 7$ . The other parameters are as indicated in the figure: the top row (A and B) has  $\delta = 0.6$  while the bottom row (C and D) has  $\delta = 1.0$ ; the left column (A and C) has super-efficient ants (with  $\alpha = 6$ ) while the right column (B and D) has inefficient ants (with  $\alpha = 1$ ). **Case A and C:** The model predictions of steady state task performance frequencies agree with simulation results. **Case D:** The model predictions and simulation results disagree in all cases, but the discrepancy was expected: since conditions (5) (for pure colonies) and (8) (for mixed colonies) are violated, we do not expect the system to be in equilibrium in the first place (in particular, the stimuli keep growing). **Case B:** The model overpredicts the steady-state task performance frequencies for the tasks at which the ants are less efficient (i.e., Task 2 for A and Task 1 for B), but in all other cases the model agrees with the simulation results.

Overall, the model performs well against the simulations when all the ants in the colony are sufficiently efficient (Fig. 1, left column). However, when some ants are inefficient (right column), the model overestimates the activity for the less efficient line, *even when the system reaches a steady state* (Case B in Fig. 1). Potential causes of this discrepancy might be:

- We assumed that the threshold  $\mu$  is constant for every genetic line and task, i.e., that there is no variation ( $\sigma = 0$ ). When I set  $\sigma = 0$  in one set of simulations, however, the discrepancy did not go away. Plus we know from previous results that  $\sigma$  only affects the variance in task performance, not the mean.
- The analytical model does not explicitly take into account the fact that, in the simulations, an ant must be inactive for one time step before it can start a task. This seems more likely, but I'm not yet sure how to incorporate this into the model.
- What else is missing??

### 3.3 Mixed colonies: 2 tasks, 2 lines, non-50-50 mixes (NEW)

We now generalize to the case in which a fraction  $f$  of individuals in a mixed colony are of genetic line **A**. In the simplified case where  $\mu_1^A = \mu_2^A = \mu_1^B = \mu_2^B$  and  $\tau^A = \tau^B$ , the steady-state fractions of individuals performing task  $j$  are

$$x_j^A = \frac{fn\delta_j}{f\alpha_j^A + (1-f)\alpha_j^B}, \quad x_j^B = \frac{(1-f)n\delta_j}{f\alpha_j^A + (1-f)\alpha_j^B}, \quad (9)$$

or, as fractions of the individuals of each genetic type (recall that there are  $fn$  individuals of type **A** and  $(1-f)n$  individuals of type **B**),

$$\frac{x_j^A}{fn} = \frac{x_j^B}{(1-f)n} = \frac{\delta_j}{f\alpha_j^A + (1-f)\alpha_j^B} \left( = \frac{x_j^A + x_j^B}{n} \right). \quad (10)$$

The last equality highlights the fact that the fraction of individuals of *each type* performing task  $j$  is identical to the fraction of *the whole colony* performing that task. As expected, the expressions (10) reduce to (7) when  $f = 0.5$  (50-50 mixes) and to (4) when  $f = 1$  (pure **A** colonies).

Again, we would expect to see this equilibria only when  $x_1^A + x_1^B + x_2^A + x_2^B \leq n$  is satisfied.

From these results, we would expect the steady-state fraction of task  $j$  performance frequency (which corresponds to (10)) to depend **nonlinearly** on the fraction of **A** individuals ( $f$ ).

#### 3.3.1 Comparing with simulation results

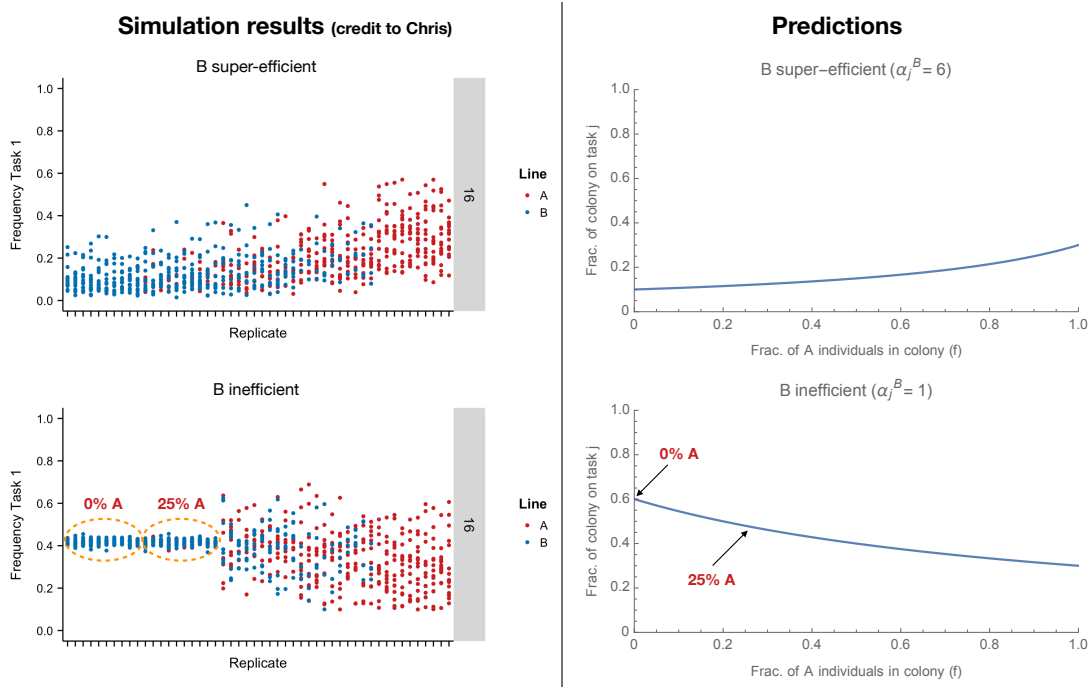


Figure 2: Comparison of simulation results (left) and analytical predictions of steady states (right) according to the model above. The parameters used in both columns are as follows:  $\alpha_1^A = \alpha_2^A = 2, \delta = 0.6, \mu = 10, \sigma = 0.1, \eta = 7$ . **Top row:** The **B** ants are “super-efficient” at both tasks ( $\alpha_1^B = \alpha_2^B = 2$ , top row). In the simulations, task performance *increases nonlinearly* as the fraction of **A** ants in the colony increases (left), as predicted qualitatively by the model (right). **Bottom row:** The **B** ants are “inefficient” at both tasks ( $\alpha_1^B = \alpha_2^B = 1$ ). Similar to the 50-50 cases, when the colony primarily consists of the inefficient **B** ants ( $f = 0, 0.25$  in the simulations), the model *overpredicts* the colony activity. Where there are sufficiently many **A** ants in the colony ( $f = 0.5, 0.75, 1$  in the simulations), however, both the simulations and model predictions show a *nonlinear negative* relationship in task performance.

## 4 Downward vs. upward contagion in the 50-50 mixes

A focus during our last Skype with Daniel and Yuko was how the model could qualitatively capture the two contagion patterns in Fig. 1 (bottom row) of Yuko’s document. Specifically, they were interested in **whether varying both the task efficiencies  $\alpha$  and the demand rates  $\delta$  would capture the difference in the directions of contagion (downward or upward) when the larvae differed.**

Here’s my conclusion so far: *If we assume that the system reaches its steady state within the observed period (i.e., conditions (5) and (8) are satisfied<sup>4</sup>), then varying **only**  $\delta$  and  $\alpha$  can result in a downward contagion but not an upward contagion.*

Here’s why I think this is the case: If we only vary  $\delta$  and  $\alpha$ , then the mean thresholds and the quit probabilities are identical across tasks and lines. So we can directly apply the steady-state fractions of active individuals computed in Eqs. (4) and (7). In this case, the behavioral contagion is “downward” if

$$\frac{1}{2} \left( \frac{\delta_j}{\alpha_j^A} + \frac{\delta_j}{\alpha_j^B} \right) > \frac{2\delta_j}{\alpha_j^A + \alpha_j^B} \quad (11)$$

and “upward” if the inequality is reversed (the schematic below should help).

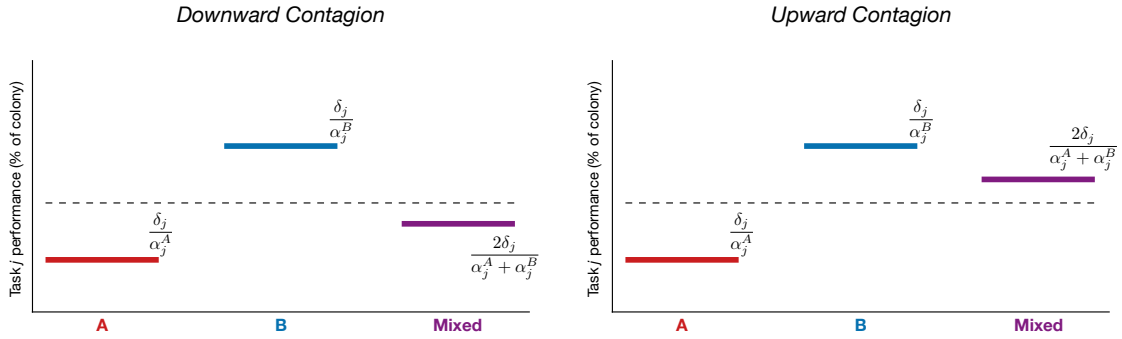


Figure 3: A schematic representing the two contagion patterns of our interest. The task  $j$  performance (%) values for the mixed colonies assume that the mean thresholds and the quit probabilities are identical for both lines and both tasks ( $\mu_1^A = \mu_2^A = \mu_1^B = \mu_2^B$ ) and ( $\tau^A = \tau^B$ ).

By manipulating the inequality (11), however, we see that **only the downward contagion is possible** under our assumptions:

$$\begin{aligned} \frac{1}{2} \left( \frac{\delta_j}{\alpha_j^A} + \frac{\delta_j}{\alpha_j^B} \right) - \frac{2\delta_j}{\alpha_j^A + \alpha_j^B} &= \frac{\delta_j}{2} \left( \frac{1}{\alpha_j^A} + \frac{1}{\alpha_j^B} - \frac{4}{\alpha_j^A + \alpha_j^B} \right) \\ &= \frac{\delta_j}{2} \left( \frac{\alpha_j^B(\alpha_j^A + \alpha_j^B) + \alpha_j^A(\alpha_j^A + \alpha_j^B) - 4\alpha_j^A\alpha_j^B}{\alpha_j^A\alpha_j^B(\alpha_j^A + \alpha_j^B)} \right) \\ &= \frac{\delta_j}{2} \left( \frac{(\alpha_j^A)^2 + (\alpha_j^B)^2 - 2\alpha_j^A\alpha_j^B}{\alpha_j^A\alpha_j^B(\alpha_j^A + \alpha_j^B)} \right) = \frac{\delta_j}{2} \left( \frac{(\alpha_j^A - \alpha_j^B)^2}{\alpha_j^A\alpha_j^B(\alpha_j^A + \alpha_j^B)} \right) > 0 \end{aligned} \quad (12)$$

Upward contagion is still possible if we relax some of our assumptions. In the simulations I’ve run so far, upward contagion appears

- when the mean thresholds ( $\mu$ ) are varied in addition to the task efficiencies ( $\alpha$ ), or
- when at least one of the lines is not sufficiently efficient to maintain the stimuli at constant levels (i.e., reach steady state).

<sup>4</sup>This part is technically not tight, since I haven’t analyzed what the sufficient conditions are

## 5 Questions

- Even when there is ample demand, the ants spend ~90% of the time performing the tasks. What mechanism explains the presence of this “ceiling” effect?
- How reasonable is it to assume that the empirically observed behavior corresponds to the steady state of the model? Need to make this assumption for steady-state analysis, but should also consider cases where this assumption does not hold and
- From our earlier simulations, we know that both the mean threshold ( $\mu$ ) or the quit probability ( $\tau$ ) lead to “behavioral amplification” in the mixed case. Biologically speaking, which is more likely to vary by genetic line? Also, what happens when  $\alpha$  AND either  $\mu$  and  $\tau$  are varied?
- What do these “upward” and “downward” contagions imply? As in, why might they be biologically interesting? We will ask Yuko et al. when we send them the summary document.

## References

- [1] Y. Ulrich, J. Saragosti, C. K. Tokita, C. E. Tarnita, D. J. C. Kronauer, “Fitness benefits and emergent division of labour at the onset of group living,” *Nature*, vol. 560, pp. 635-638, Aug. 2018.