eDNA

Richard Arnold

2024-07-16

Detection of eDNA

- A volume V_S of water is collected, and all of the eDNA it contains is collected by filtration. Assume a count N_j of eDNA fragments of type j is collected ($j = 1, \dots, J$). If the number density of fragments in the water is ν_j then
$$N_j|\nu_j, V_S \sim \text{Poisson}(\nu_j V_S)$$
- The collected eDNA is mixed with a volume V_M of a proprietary mix of primers, probes and water;
- A test volume V_T ($1\mu\ell$) is taken from the mixture: The count of fragments in the test volume is
$$n_j|N_j, V_M, V_T \sim \text{Poisson}(N_j V_T/V_M)$$
- The test volume is atomised into m droplets. The count of fragments in droplet i is
$$y_{ji}|n_j, m \sim \text{Poisson}(n_j/m)$$
- The probability that droplet i has zero fragments is
$$\Pr(y_{ji} = 0|n_j, m) = \exp\left(-\frac{n_j}{m}\right)$$
and thus the probability that it contains any fragments is
$$\Pr(y_{ji} > 0|n_j, m) = 1 - \exp\left(-\frac{n_j}{m}\right)$$
- The eDNA in the droplets is amplified by PCR, and m_{j+} of the droplets are observed to be positive for fragment type j. Thus the observed proportion of positive droplets is
$$\widehat{p}_+ = \frac{m_{j+}}{m}$$
and
$$m_{j+}|n_j, m \sim \text{Binomial}\left(m, 1 - e^{-n_j/m}\right)$$
or if n_j/m is very small then
$$m_{j+}|n_j, m \sim \text{Poisson}\left(m\left(1 - e^{-n_j/m}\right)\right)$$
Note that we assume no false positives and no false negatives when individual droplets are amplified and tested for the presence of the eDNA.

This process is aimed at determining ν_j, the original number density in the water from which the sample was collected. A simple estimate of μ_j is

$$\widehat{\nu}_j = -\frac{V_M m}{V_S V_T} \log\left(1 - \frac{m_{j+}}{m}\right)$$

However in this application we are most interested in the probability that the process leads to negative result when ν_j is non-zero. i.e. we want to know

$$
\begin{aligned}
& \Pr(m_{j+}=0|m,V_S,V_T,V_M,\nu_j) && (1)\\
&= \sum_{N_j=0}^{\infty}\sum_{n_j=0}^{\infty} \Pr(m_{j+}=0|n_j,m)\Pr(n_j|N_j,V_M,V_T)\Pr(N_j|\nu_j,V_S) && (2)\\
&= \sum_{N_j=0}^{\infty}\sum_{n_j=0}^{\infty} \exp\left(-m\left[1-e^{-n_j/m}\right]\right)\exp\left(-\frac{N_jV_T}{V_M}\right)\frac{\left(\frac{N_jV_T}{V_M}\right)^{n_j}}{n_j!}\exp\left(-\nu_jV_S\right)\frac{(\nu_jV_S)^{N_j}}{N_j!} && (3)
\end{aligned}
$$

Note that the sum over n_j should only be from 0 to N_j, but we sum to infinity because the additional terms are all negligible.

Writing the ratio V_T/V_M, i.e. the proportion of the mixture that is atomised and tested, as d, and suppressing the index j

$$
\begin{aligned}
& \Pr(m_+=0|m,V_S,d,\nu) && (4)\\
&= \sum_{N=0}^{\infty}\sum_{n=0}^{\infty} \Pr(m_+=0|n,m)\Pr(n|N,d)\Pr(N|\nu,V_S) && (5)\\
&= \sum_{N=0}^{\infty}\sum_{n=0}^{\infty} \exp\left(-m\left[1-e^{-n/m}\right]\right)e^{-Nd}\frac{(Nd)^n}{n!}e^{-\nu V_S}\frac{(\nu V_S)^N}{N!} && (6)
\end{aligned}
$$

At very low concentrations

$$n|\nu,V_S,d \sim \text{Poisson}(\nu V_S d)$$

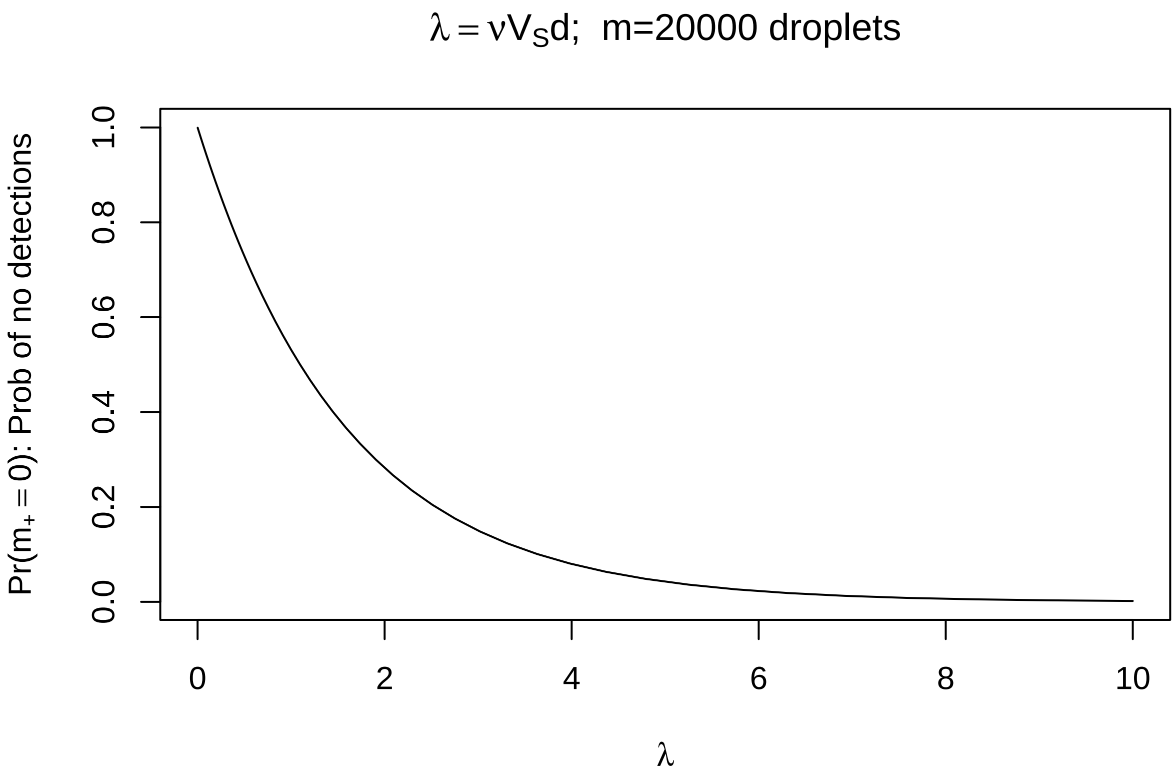
so that

$$
\begin{aligned}
& \Pr(m_+=0|m,V_S,d,\nu) && (7)\\
&= \sum_{n=0}^{\infty} \Pr(m_+=0|n,m)\Pr(n|\nu,V_S,d) && (8)\\
&= \sum_{n=0}^{\infty} \exp\left(-m\left[1-e^{-n/m}\right]\right)e^{-\nu V_S d}\frac{(\nu V_S d)^n}{n!} && (9)\\
&= \sum_{n=0}^{\infty} \exp\left(-m\left[1-e^{-n/m}\right]\right)e^{-\lambda}\frac{\lambda^n}{n!} && (10)\\
& && (11)
\end{aligned}
$$

where $\lambda=\nu V_S d$.

Typically we have $m=2\times10^4$ droplets. We might expect $\nu V_s \simeq 10$ fragments to be catpured, and a subsampling of $d=1/20$, and in that case $\lambda \simeq 0.5$

```
probmp0 <- function(lambda, m=20000, nbig=20000) {
  nvec <- 0:nbig
  fval <- sapply(lambda, function(lambdav) sum(exp(-m*(1-exp(-nvec/m)))*dpois(nvec,lambdav)))
  return(fval)
}
m <- 20000; nbig <- m
np <- 101
lambdavec <- 10^(seq(from=-3,to=+1,length=np))
plot(lambdavec, (probmp0(lambdavec, m=m, nbig=nbig)), type="l",
     xlab=expression(lambda), ylab=expression("Pr("*m["+"]==0*"): Prob of no detections"),
     main=bquote(lambda==nu*V[S]*d*";  "*m*"="*.(m)*" droplets"))
```



Particle dispersal

If there is a mass released at time t_0 at location $\mathbf{x}_0$ then it is transported by advection (water flow) and dispersion (turbulence).

Let $p(\mathbf{x}, t|\mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)$ be the proportion of the mass released that is found at location $\mathbf{x}$ at time t, given water flow $\mathbf{v}()$ and dispersion $\mathbf{a}()$. Ψ is a set of additional parameters controlling the flow/dispersion model.

A simple model is a simple Gaussian spread in 2D:

$$
\begin{aligned}
p(\mathbf{x}, t|\mathbf{x}_0, t_0, \mathbf{v}, \mathbf{a}, \Psi) &= \frac{I(t - t_0 \geq 0)}{2\pi\sqrt{a_x a_y}(t - t_0 + t_\varepsilon)} \exp\left(-\frac{(x - x_0 - v_x(t - t_0))^2}{2a_x(t - t_0 + t_\varepsilon)} - \frac{(y - y_0 - v_y(t - t_0))^2}{2a_y(t - t_0 + t_\varepsilon)}\right) && (12)\\
&= g(\mathbf{x}, t - t_0|\mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) && (13)
\end{aligned}
$$

where a small positive temporal offset $\Psi = \{t_\varepsilon\}$ is added to avoid a singularity at $t = t_0$.

We have

$$\int\!\!\int p(\mathbf{x}, t|\mathbf{x}_0, t_0, \mathbf{v}, \mathbf{a}, \Psi)\, \mathrm{d}\mathbf{x} = 1$$

for all times $t > t_0$.

Note that $g()$ has a symmetry in which we can interchange $\mathbf{x}$ and $\mathbf{x}_0$ and at the same time reverse the velocity field, so for any time interval $u \geq 0$:

$$g(\mathbf{x}, u|\mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = g(\mathbf{x}_0, u|\mathbf{x}, -\mathbf{v}, \mathbf{a}, \Psi) \tag{14}$$

```r
source("funcs.R")
```

```r
nx <- 101
ny <- 101
xmin <- 0; xmax <- 100
```

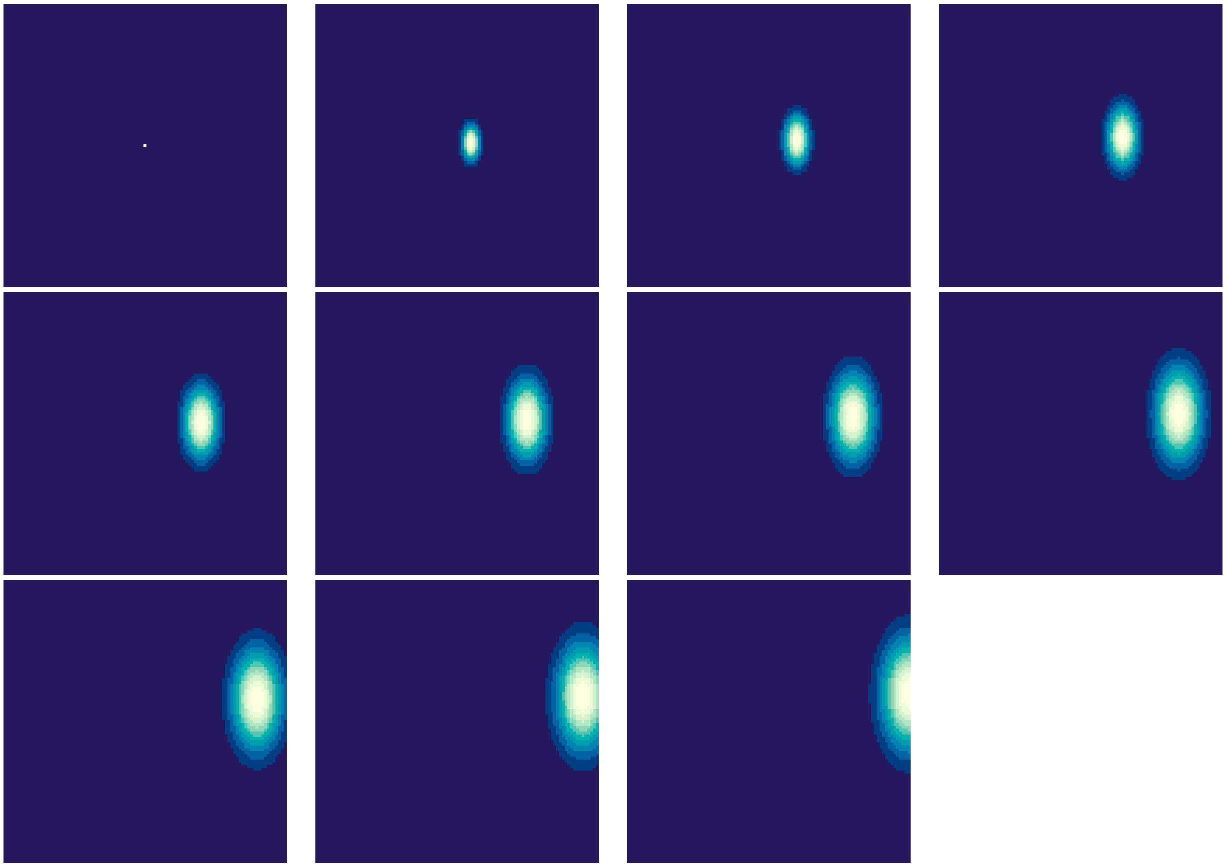
```r
ymin <- 0; ymax <- 100
xvec <- seq(from=xmin, to=xmax, length=nx)
yvec <- seq(from=ymin, to=ymax, length=ny)

x0 <- mean(c(xmin,xmax))
y0 <- mean(c(ymin,ymax))
t0 <- 0
vx <- 5; vy <- 1
ax <- 2^2; ay <- 4^2
teps <- 0.01
```

For example here is the spread over time of mass released at $\mathbf{x}_0 = (50, 50)^T$ at time $t = 0$ with constant flow field $\mathbf{v} = (5, 1)^T$ and dispersion $\mathbf{a} = (4, 16)^T$ and with $t_\varepsilon = 0.01$.

Forward tracking: $p(\mathbf{x}, t|\mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)$: $t_0, \ldots, t_0 + 10$ (no decay)

```r
par(mfrow=c(3,4))
par(mar=0.1*c(1,1,1,1))
colvec <- hcl.colors(12, "YlOrRd", rev = TRUE)
colvec <- hcl.colors(12, "YlGnBu")#, rev = TRUE)
for(t in 0:10) {
   amat <- outer(xvec,yvec, dispfunc, t=t,
                 x0=x0,y0=y0,t0=0, vx=vx,vy=vy, ax=ax,ay=ay, teps=teps)
   #if(t==0) amax <- max(amat)
   amax <- max(amat)
   breaks <- seq(from=0, to=amax, length=length(colvec)+1)
   image(xvec, yvec, amat, asp=1, col=colvec, breaks=breaks, axes=FALSE)
   #image(xvec, yvec, amat,asp=1, main=bquote("Forwards:" ~ t==.(t)), xlab="x", ylab="y",
   #      col=colvec, breaks=breaks)
}
```



A simple backtracking model simply reverses the flow field. If we are interested in the likely origin of mass observed at location $\mathbf{x}_1$ at time t_1 we can backtrack by reversing $\mathbf{v}()$.

We might assume that the likelihood that the observed mass was released at location $\mathbf{x}$ at time $t < t_1$ is given by

$$\begin{aligned} \tilde{p}(\mathbf{x}, t|\mathbf{x}_1, t_1, \mathbf{v}(), \mathbf{a}(), \Psi) &= g(\mathbf{x}, t_1 - t|\mathbf{x}_1, -\mathbf{v}(), \mathbf{a}(), \Psi) && (15) \\ &= p(\mathbf{x}, -t|\mathbf{x}_1, -t_1, -\mathbf{v}(), \mathbf{a}(), \Psi) && (16) \end{aligned}$$

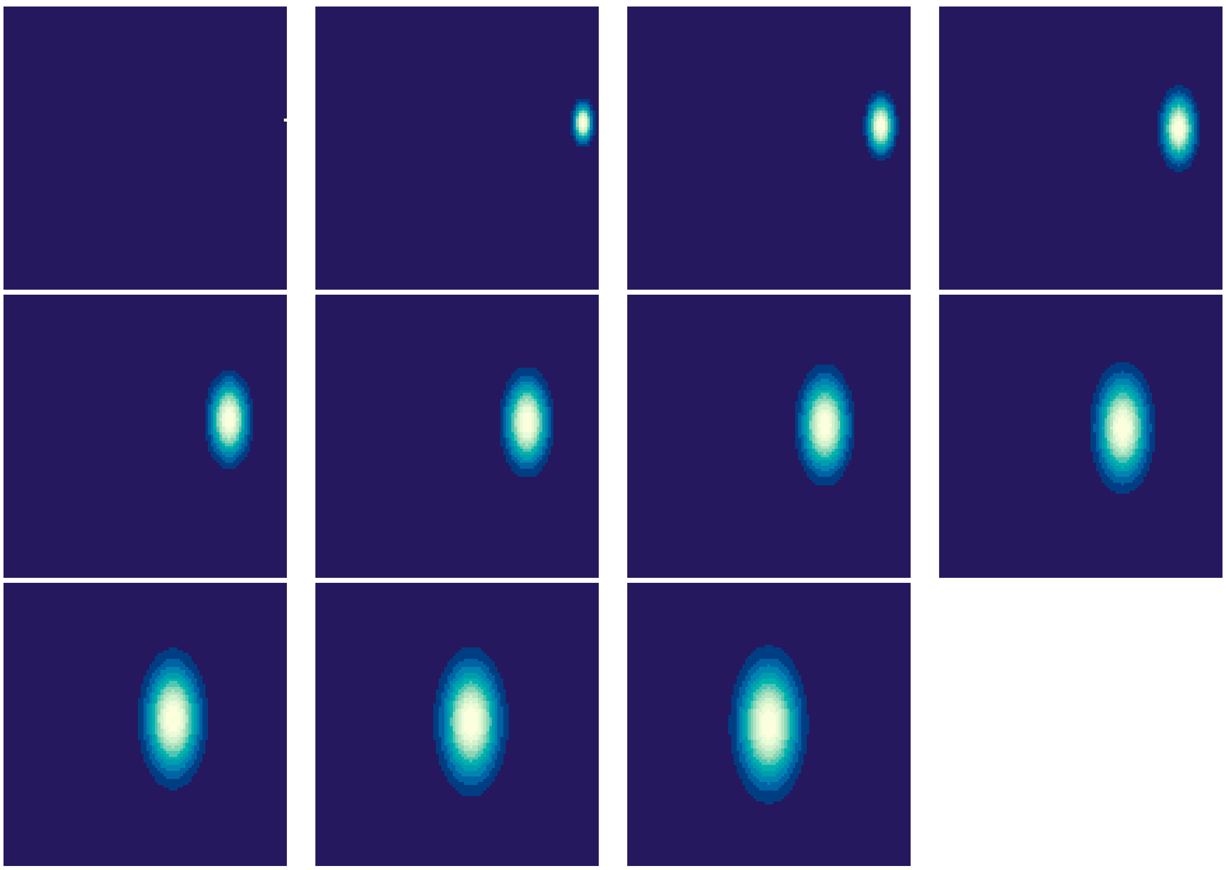
In the simple Gaussian model above this is

$$\begin{aligned} &\tilde{p}(\mathbf{x}, t|\mathbf{x}_1, t_1, \mathbf{v}(), \mathbf{a}(), \Psi) && (17) \\ &\quad = p(\mathbf{x}, t_1|\mathbf{x}_1, t, -\mathbf{v}(), \mathbf{a}(), \Psi) && (18) \\ &\quad = p(\mathbf{x}, -t|\mathbf{x}_1, -t_1, -\mathbf{v}(), \mathbf{a}(), \Psi) && (19) \\ &\quad = g(\mathbf{x}, t_1 - t|\mathbf{x}_1, -\mathbf{v}(), \mathbf{a}(), \Psi) && (20) \\ &\quad = \frac{I(t_1 - t \geq 0)}{2\pi\sqrt{a_x a_y}(t_1 - t + t_\varepsilon)} \exp\left(-\frac{(x - x_1 + v_x(t_1 - t))^2}{2a_x(t_1 - t + t_\varepsilon)} - \frac{(y - y_1 + v_y(t_1 - t))^2}{2a_y(t_1 - t + t_\varepsilon)}\right) && (21) \end{aligned}$$

Back tracking: $\tilde{p}(\mathbf{x}, t|\mathbf{x}_1, t_1, \mathbf{v}(), \mathbf{a}(), \Psi) = p(\mathbf{x}, -t|\mathbf{x}_1, -t_1, -\mathbf{v}(), \mathbf{a}(), \Psi)$: $t_1, \ldots, t_1 - 10$ (no decay)

```
par(mfrow=c(3,4))
par(mar=0.1*c(1,1,1,1))
x1 <- 100; y1 <- 60
t1 <- 10
for(t in 10:0) {
  amat <- outer(xvec,yvec, dispfunc, t=-t,
```

```
                 x0=x1,y0=y1,t0=-t1, vx=-vx,vy=-vy, ax=ax,ay=ay, teps=teps)
  #if(t==10) amax <- max(amat)
  amax <- max(amat)
  breaks <- seq(from=0, to=amax, length=length(colvec)+1)
  image(xvec, yvec, amat,asp=1, col=colvec, breaks=breaks, axes=FALSE)
  #image(xvec, yvec, amat,asp=1, main=bquote("Backwards: " ~ t==.(t)), xlab="x", ylab="y",
  #       col=colvec, breaks=breaks)
}
```



Continuous release with decay

Assume that at location $\mathbf{x}_0$ there is a continuous release of mass at a rate N_0 kg/s. Further assume that the mass decays at a rate

$$d(t|T_h, T_m) = 2^{-t/T_h} I(0 \leq t \leq T_m)$$

where T_h is the decay half life and T_m is the maximum age at which all particles have disintegrated. If we set $\lambda = (1/T_h)\log 2$ then

$$d(t|T_h, T_m) = e^{-\lambda t} I(0 \leq t \leq T_m)$$

The steady state mass in the system is then

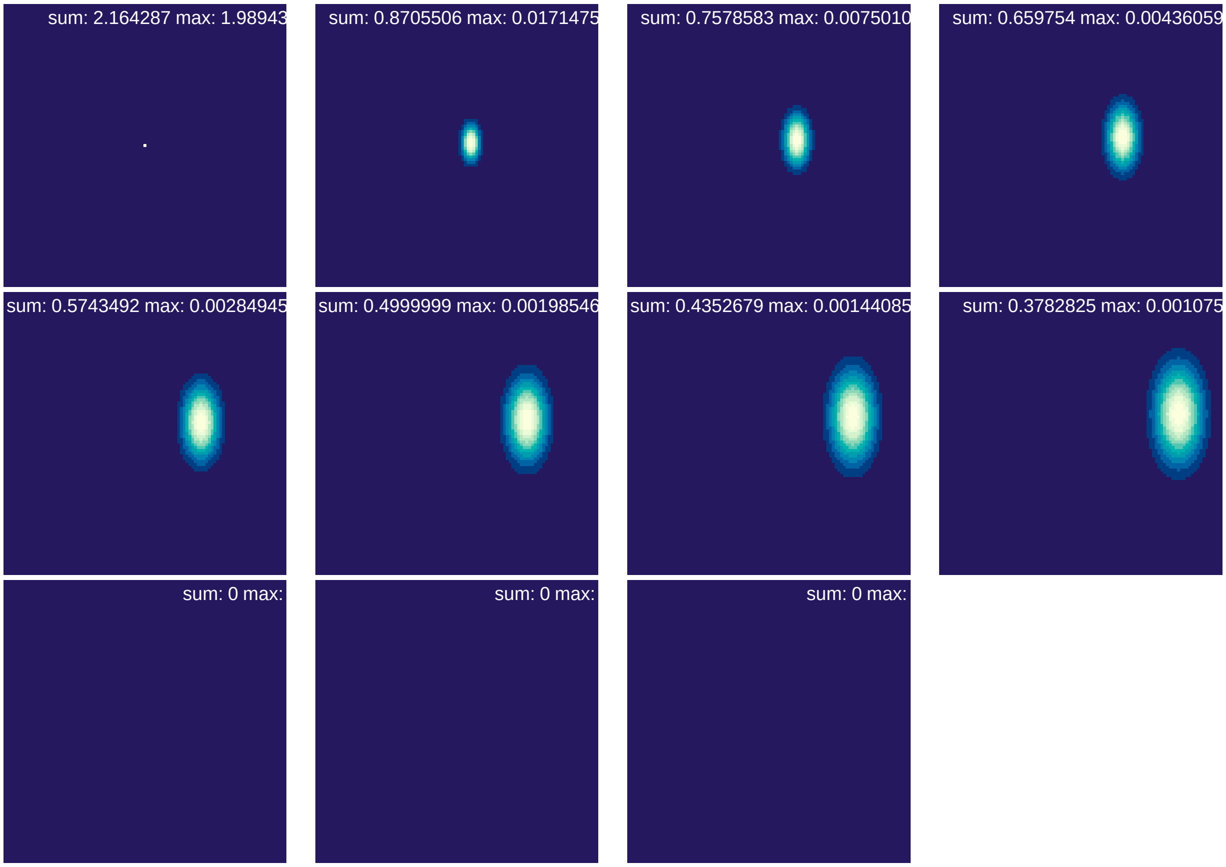
$$
\begin{aligned}
M_0 &= \int_{-\infty}^{t} N_0 d(t - t_0 | T_h, T_m)\ \mathrm{d}t_0 && (22)\\
&= N_0 \int_{-\infty}^{t} e^{-\lambda(t-t_0)} I(t_0 > t - T_m)\ \mathrm{d}t_0 && (23)\\
&= N_0 \int_{t-T_m}^{t} e^{-\lambda(t-t_0)}\ \mathrm{d}t_0 && (24)\\
&= N_0 \int_{0}^{T_m} e^{-\lambda u}\ \mathrm{d}u && (25)\\
&= \frac{N_0}{\lambda}\left[1 - e^{-\lambda T_m}\right] && (26)\\
&= \frac{N_0 T_h}{\log 2}\left[1 - e^{-\lambda T_m}\right] && (27)
\end{aligned}
$$

The mass density transiting per unit time at location $\mathbf{x}$ at time t emitted from a point source at $\mathbf{x}_0$ at time t_0 is

$$\frac{\mathrm{d}\rho(\mathbf{x}, t | N_0, \mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)}{\mathrm{d}t} = N_0 d(t - t_0 | T_h, T_m) p(\mathbf{x}, t | \mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)$$

Density rate $\mathrm{d}\rho(\mathbf{x}, t | N_0, \mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)/\mathrm{d}t$, $t = t_0, \ldots, t_0 + 10$ (with decay). (Scales differ between panels: note max and sum values.)

```
par(mfrow=c(3,4))
par(mar=0.1*c(1,1,1,1))
n0 <- 1
th <- 5
tm <- 7
for(t in 0:10) {
   amat <- outer(xvec,yvec, denratefunc, t=t,
                 x0=x0,y0=y0,t0=0, vx=vx,vy=vy, ax=ax,ay=ay, teps=teps, n0=n0,th=th,tm=tm)
   #amat <- log(1+amat)
   #if(t==0) {
   #   amin <- min(amat)
   #   amax <- max(amat)
   #}
   amin <- min(amat)
   amax <- max(amat)
   breaks <- seq(from=amin, to=amax, length=length(colvec)+1)
   image(xvec, yvec, amat, asp=1, col=colvec, breaks=breaks, axes=FALSE)
   #image(xvec, yvec, amat,asp=1, main=bquote("Forwards with decay:" ~ t==.(t)), xlab="x", ylab="y",
   #       col=colvec, breaks=breaks)
   mtext(bquote("sum:"~.(sum(amat))~"max:"~.(max(amat))), side=3, adj=1, cex=0.6, line=-1, col="white")
}
```



The total (equilibrium) mass density present is

$$
\begin{aligned}
&\rho(\mathbf{x}, t|N_0, \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) && (28)\\
&\quad = \rho(\mathbf{x}|N_0, \mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) && (29)\\
&\quad = \int_{-\infty}^{t} N_0 d(t - t_0|T_h, T_m) p(\mathbf{x}, t|\mathbf{x}_0, t_0, \mathbf{v}(), \mathbf{a}(), \Psi)\, \mathrm{d}t_0 && (30)\\
&\quad = N_0 \int_{t-T_m}^{t} e^{-\lambda(t-t_0)} g(\mathbf{x}, t - t_0|\mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi)\, \mathrm{d}t_0 && (31)\\
&\quad = N_0 \int_{0}^{T_m} e^{-\lambda u} g(\mathbf{x}, u|\mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi)\, \mathrm{d}u && (32)\\
&\quad = N_0 h(\mathbf{x}|\mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) && (33)
\end{aligned}
$$

where

$$h(\mathbf{x}|\mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) = \int_0^{T_m} e^{-\lambda u} g(\mathbf{x}, u|\mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi)\, \mathrm{d}u$$

The function $h()$ retains the same symmetry as $g()$, so since

$$g(\mathbf{x}, u|\mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = g(\mathbf{x}_0, u|\mathbf{x}, -\mathbf{v}, \mathbf{a}, \Psi)$$

it follows that

$$h(\mathbf{x}, u|\mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = h(\mathbf{x}_0, u|\mathbf{x}, -\mathbf{v}, \mathbf{a}, \Psi)$$

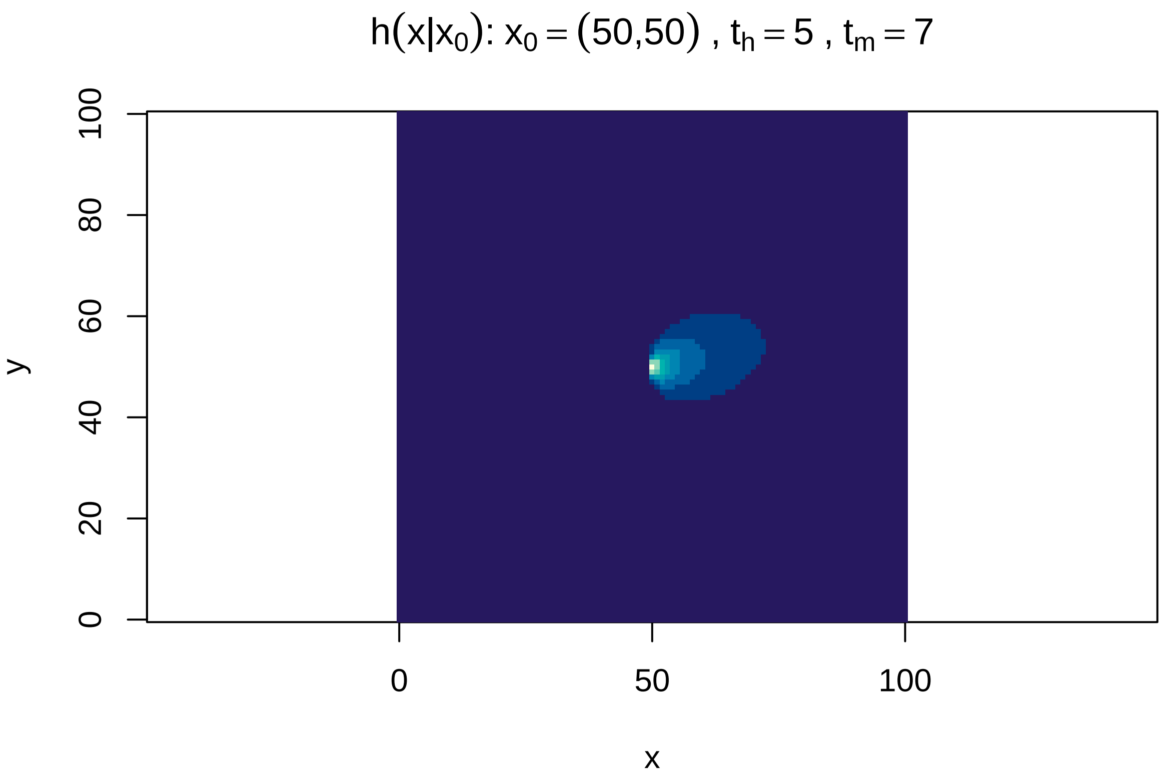
In the Gaussian dispersal example

$$h(\mathbf{x}|\mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi) = \frac{1}{2\pi\sqrt{a_x a_y}} \int_0^{T_m} \frac{e^{-\lambda u}}{u + t_\varepsilon} \exp\left(-\frac{(x - x_0 - v_x u)^2}{2a_x(u + t_\varepsilon)} - \frac{(y - y_0 - v_y u)^2}{2a_y(u+, t_\varepsilon)}\right)$$

```r
invisible()
#th <- 5
#tm <- 7
#ax <- 2^2; ay <- 5^2
#amat <- array(0,dim=c(nx,ny))
#nc <- 100
#for(i in 1:nc) {
#   t <- i/nc * 10
#   amat <- amat + outer(xvec,yvec, denratefunc, t=t,
#                          x0=x0,y0=y0,t0=0, vx=vx,vy=vy, ax=ax,ay=ay, teps=teps, n0=n0,th=th,tm=tm)
#}
#image(xvec, yvec, amat,asp=1,
#      main=bquote(t[h]==.(th) ~ ", " ~ t[m]==.(tm)),
#      xlab="x", ylab="y",col=colvec)
invisible()
```

Equilibrium Density: $\rho(\mathbf{x}|N_0, \mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi) = N_0 h(\mathbf{x}|\mathbf{x}_0, \mathbf{v}, \mathbf{a}, \Psi)$

```r
th <- 5
tm <- 7
ax <- 2^2; ay <- 5^2
amat <- n0*outer(xvec,yvec, hfunc,
                 x0=x0,y0=y0, vx=vx,vy=vy, ax=ax,ay=ay, teps=teps, th=th,tm=tm)
image(xvec, yvec, amat, asp=1,
      main=bquote(h(x*"|"*x[0])*":" ~ x[0]==(.(x0)*","*.(y0)) ~ "," ~ t[h]==.(th) ~ "," ~ t[m]==.(tm)),
```



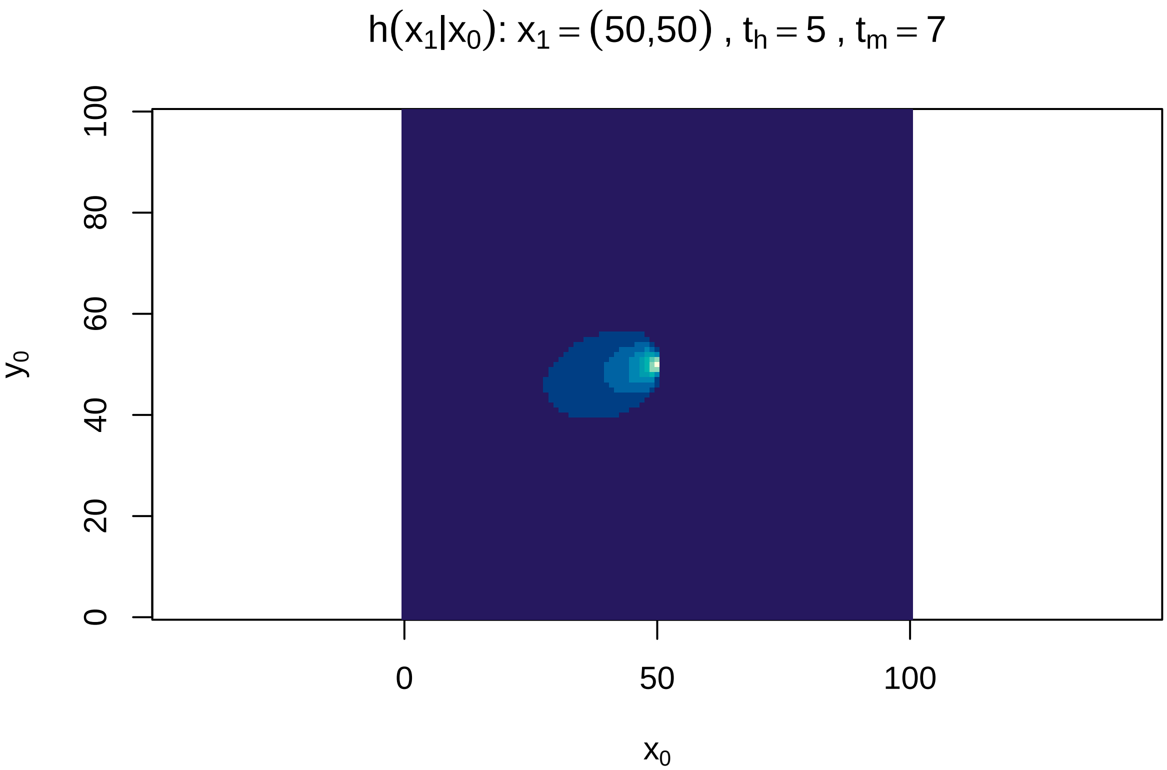
The above figure plots, for a continuous fixed source of rate N_0 at location $\mathbf{x}_0$, the equilibrium density at all locations $\mathbf{x}$: $N_0 h(\mathbf{x}|\mathbf{x}_0)$.

Now plot at each $\mathbf{x}_0$ the equilibrium density at $\mathbf{x}_1$ (in the centre of the diagram) that would result from a continuous source at $\mathbf{x}_0$. i.e. $N_0 h(\mathbf{x}_1|\mathbf{x_0})$

$$\begin{aligned} N_0\tilde{h}(\mathbf{x}_0|\mathbf{x}_1,\mathbf{v},\mathbf{a},\Psi) &= N_0 h(\mathbf{x}_1|\mathbf{x}_0,\mathbf{v},\mathbf{a},\Psi) \\ &= N_0 h(\mathbf{x}_0|\mathbf{x}_1,-\mathbf{v},\mathbf{a},\Psi) \end{aligned}$$

$N_0\tilde{h}(\mathbf{x}_0|\mathbf{x}_!,\mathbf{v},\mathbf{a},\Psi) = N_0 h(\mathbf{x}_1|\mathbf{x}_0,\mathbf{v},\mathbf{a},\Psi) = N_0 h(\mathbf{x}_0|\mathbf{x}_1,-\mathbf{v},\mathbf{a},\Psi)$

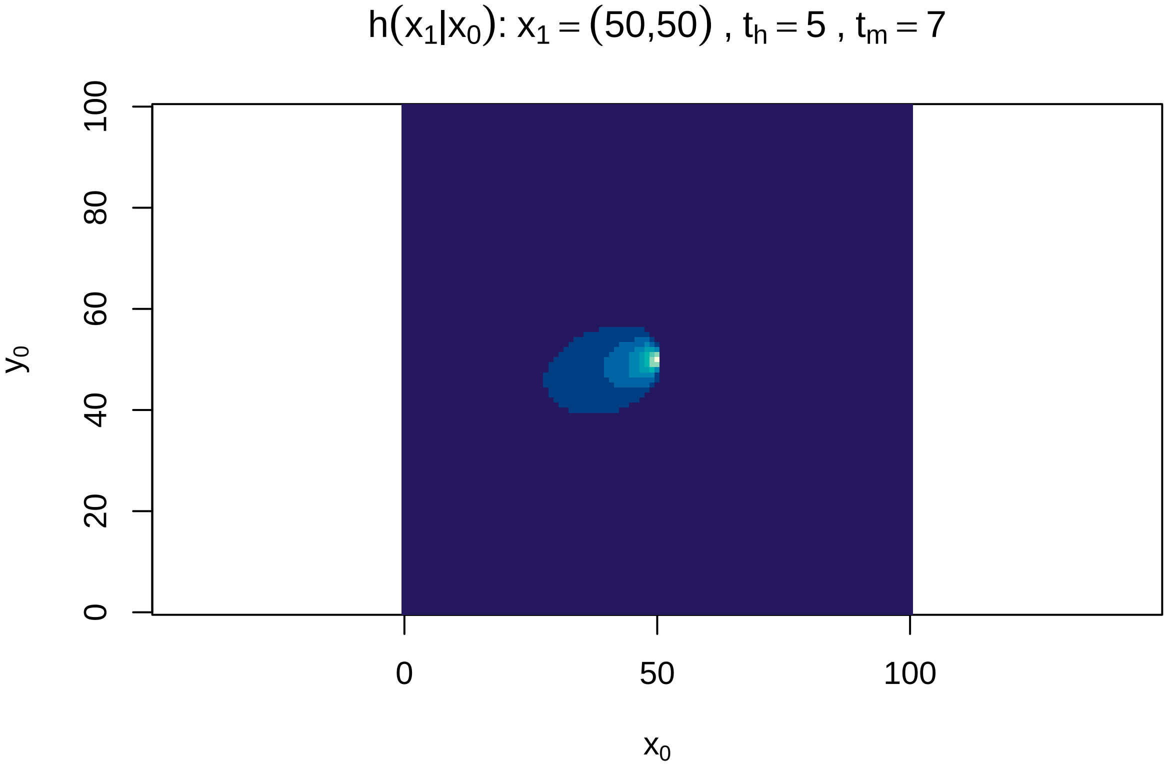
```
th <- 5
tm <- 7
ax <- 2^2; ay <- 5^2
x1 <- 50; y1 <- 50
rho1 <- 1; mu <- 1
amat <- outer(xvec,yvec, function(x0vec,y0vec) {
                    apply(cbind(x0vec,y0vec),1,function(x0y0) {
                          hfunc(x=x1,y=y1,x0=x0y0[1],y0=x0y0[2],
                          vx=vx,vy=vy, ax=ax,ay=ay, teps=teps, th=th,tm=tm)
                    })
                 })
dim(amat) <- c(length(xvec),length(yvec))
image(xvec, yvec, amat, asp=1,
      main=bquote(h(x[1]*"|"*x[0])*":" ~ x[1]==(.(x0)*","*.(y0)) ~ "," ~ t[h]==.(th) ~ "," ~ t[m]==.(tm
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
```



(Computing this the other way using the symmetry of h())

```
th <- 5
tm <- 7
ax <- 2^2; ay <- 5^2
x1 <- 50; y1 <- 50
rho1 <- 1; mu <- 1
amat <- n0*outer(xvec,yvec, hfunc,
```

```
                   x0=x1,y0=y1, vx=-vx,vy=-vy, ax=ax,ay=ay, teps=teps, th=th,tm=tm)
image(xvec, yvec, amat, asp=1,
      main=bquote(h(x[1]*"|"*x[0])*":" ~ x[1]==(.(x0)*","*.(y0)) ~ "," ~ t[h]==.(th) ~ "," ~ t[m]==.(tm
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
```



Probabilistic Inversion

Now assume that at some location $\mathbf{x}_1$ we observe a density ρ_1, and that this in an observation made with error:

$$\rho_1|\rho_1^*, \sigma_e \sim N(\rho_1^*, \sigma_e^2)$$

were the true density is ρ_1^* and the error variance is σ_e^2.

Assume that there is a single continous source of unknown strength N_0 and unknown location $\mathbf{x}_0$. Priors for the strength and location are $\pi(N_0)$ and $\pi(\mathbf{x}_0)$ respectively.

Assume that the velocity field $\mathbf{v}()$ and the diffusion field $\mathbf{a}()$ are known.

Using the results from the previous section, conditional on N_0 and $\mathbf{x}_0$ the observed density at $\mathbf{x}_1$ is

$$\rho_1^* = N_0 h(\mathbf{x}_1|\mathbf{x}_0, \mathbf{v}(), \mathbf{a}(), \Psi)$$

which we write for brevity as

$$\rho_1^* = h_1^*(\mathbf{x}_0) = h_1^*$$

but noting its dependence on $\mathbf{x}_0$.

It follows that the joint distribution of the unknown source strength N_0 and location $\mathbf{x}_0$ is

$$\begin{aligned}
p(N_0, \mathbf{x}_0|\rho_1, \mathbf{x}_1) &\propto \pi(N_0)\pi(\mathbf{x}_0)\exp\left(-\frac{1}{2\sigma_e^2}\left[\rho_1 - N_0 h_1^*(\mathbf{x}_0)\right]^2\right) \\
&\propto \pi(N_0)\pi(\mathbf{x}_0)\exp\left(-\frac{(h_1^*)^2}{2\sigma_e^2}\left[N_0 - \frac{\rho_1}{h_1^*}\right]^2\right)
\end{aligned}$$

A suitable prior for N_0 is the Exponential(μ) distribution

$$\begin{aligned} N_0 &\sim \text{Exp}(\mu) \\ \pi(N_0) &= \mu e^{-\mu N_0} \end{aligned}$$

This can be marginalised out of the posterior distribution as follows. Firstly the joint distribution is

$$\begin{aligned} p(N_0, \mathbf{x}_0|\rho_1, \mathbf{x}_1) &\propto \pi(\mathbf{x}_0) e^{-\mu N_0} \exp\left(-\frac{(h_1^*)^2}{2\sigma_e^2}\left[N_0 - \frac{\rho_1}{h_1^*}\right]^2\right) \\ &\propto \pi(\mathbf{x}_0) \exp\left(-\frac{(h_1^*)^2}{2\sigma_e^2}\left[N_0 - \left(\frac{\rho_1}{h_1^*} - \frac{\mu\sigma_e^2}{(h_1^*)^2}\right)\right]^2 - \frac{\rho_1\mu}{h_1^*} + \frac{\mu^2\sigma_e^2}{2(h_1^*)^2}\right) \end{aligned}$$

Integrating out N_0 leads to

$$p(\mathbf{x}_0|\rho_1, \mathbf{x}_1) \propto \pi(\mathbf{x}_0)\frac{\sigma_e}{h_1^*} \exp\left(-\frac{\rho_1\mu}{h_1^*} + \frac{\mu^2\sigma_e^2}{2(h_1^*)^2}\right)$$

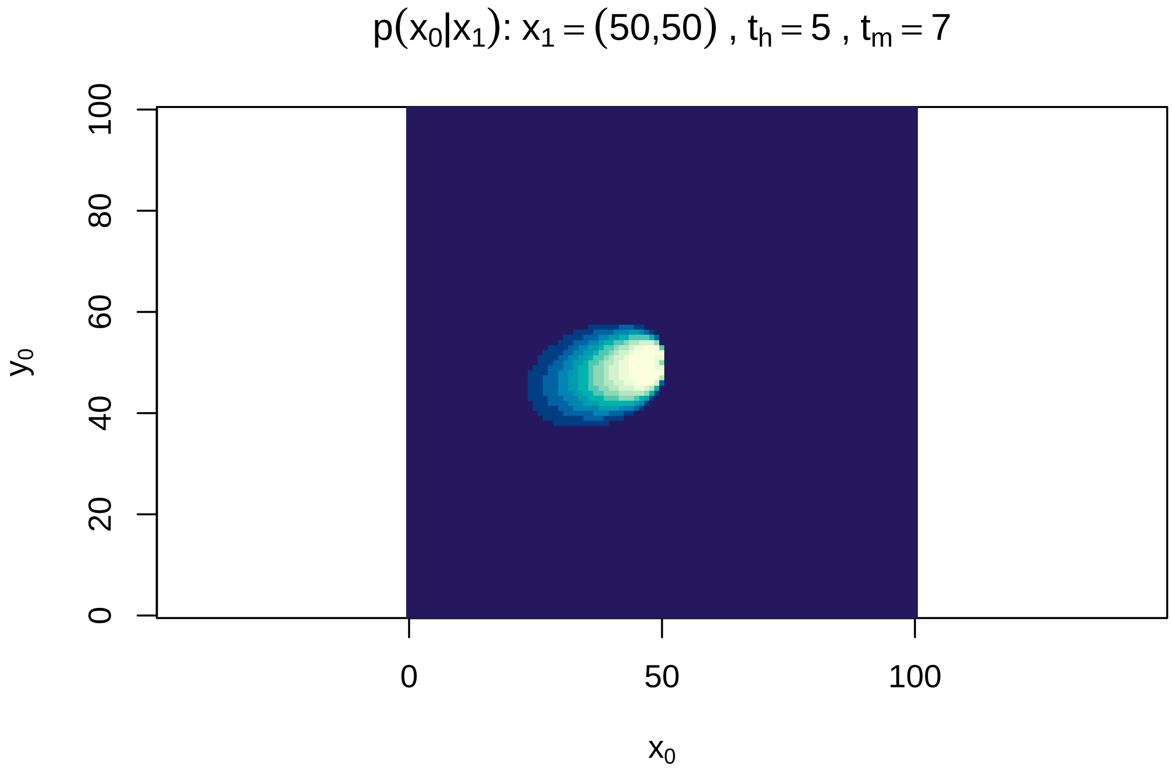
If we let $\sigma_e \to 0$ (indicating very accurate observations) we are left with a very simple form for the posterior:

$$p(\mathbf{x}_0|\rho_1, \mathbf{x}_1) \propto \pi(\mathbf{x}_0)\frac{1}{h_1^*} e^{\frac{\rho_1\mu}{h_1^*}}$$

which we plot below for a uniform prior on the source location $\pi(\mathbf{x}_0) \propto 1$.

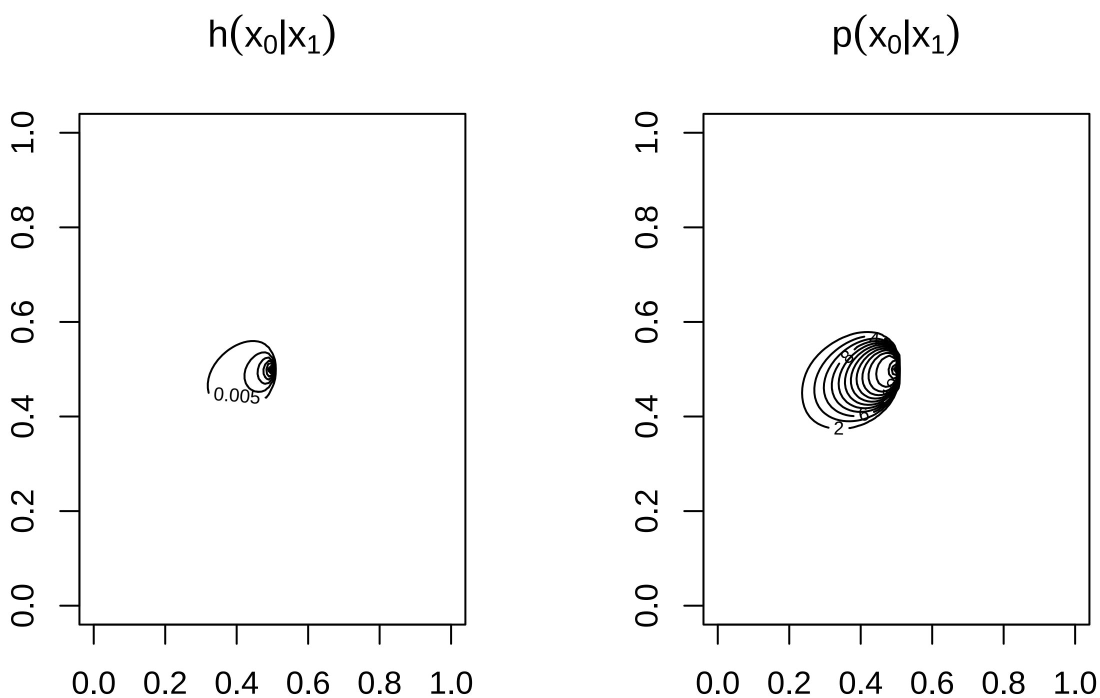
```
th <- 5
tm <- 7
ax <- 2^2; ay <- 5^2
x1 <- 50; y1 <- 50
rho1 <- 1; mu <- 1
amat <- n0*outer(xvec,yvec, hfunc,
                 x0=x1,y0=y1, vx=-vx,vy=-vy, ax=ax,ay=ay, teps=teps, th=th,tm=tm)

qmat <- exp(-max(amat)/3.0/amat) * (1/amat)
image(xvec, yvec, qmat, asp=1,
      main=bquote(p(x[0]*"|"*x[1])*":" ~ x[1]==(.(x0)*","*.(y0)) ~ "," ~ t[h]==.(th) ~ "," ~ t[m]==.(tm
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
```



Plotting again, but as a contour map

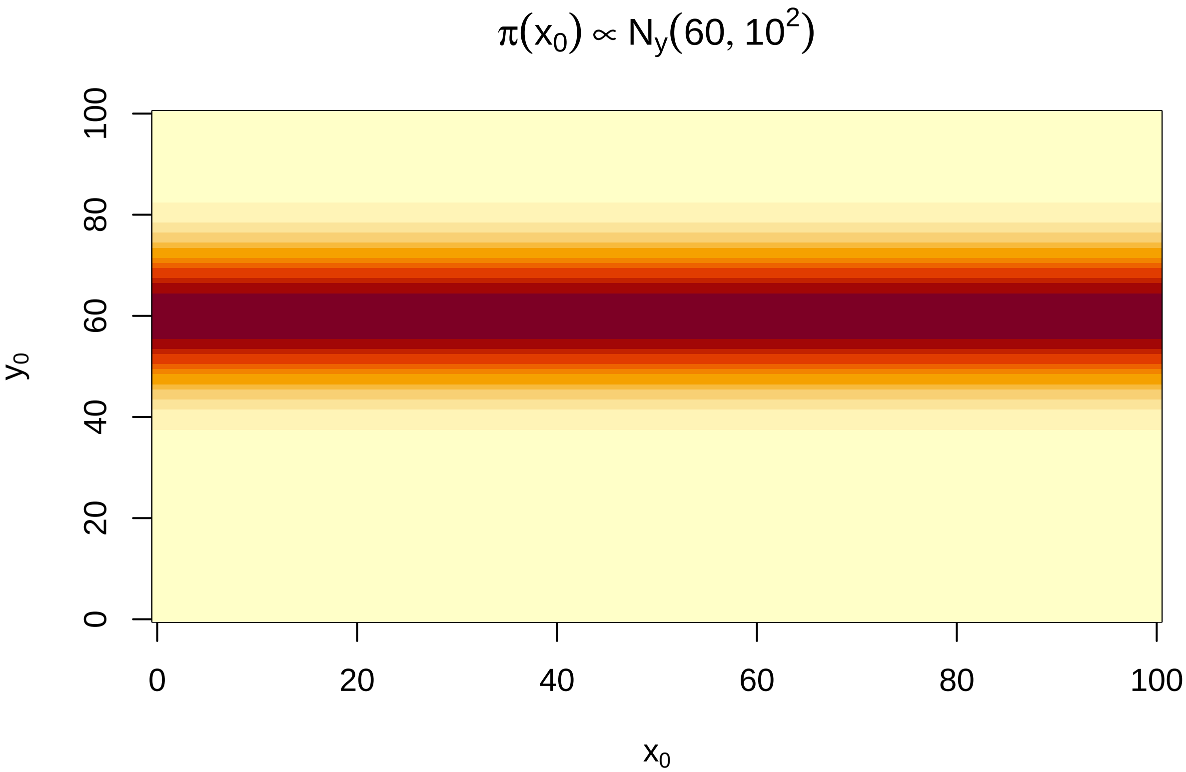
```
par(mfrow=c(1,2))
contour(amat, main=expression(h(x[0]*"|"*x[1])))
contour(qmat, main=expression(p(x[0]*"|"*x[1])))
```



A different prior: e.g. normally distributed in y (centered on y_a with standard deviation σ_a) and uniform in x:

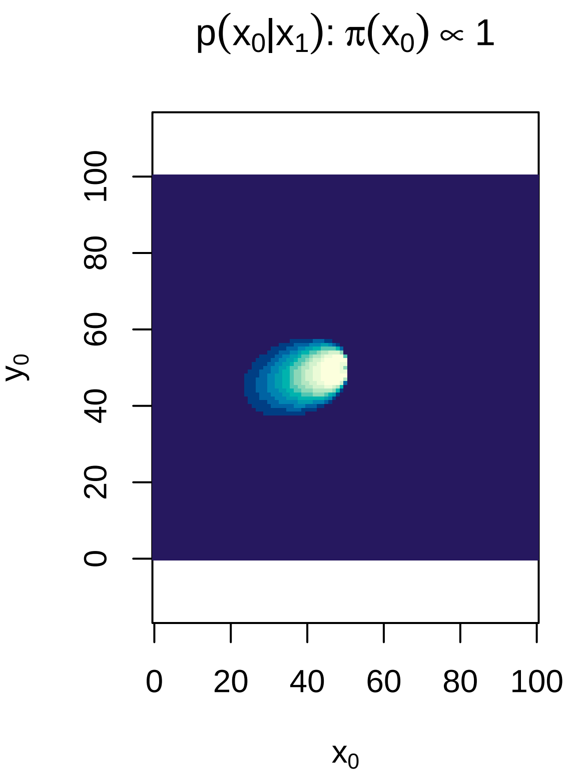
$$\pi(\mathbf{x}_0) \propto \exp\left(-\frac{1}{2\sigma_a^2}\left[x_{0y} - y_a\right]^2\right)$$

```r
ya <- 60
sigmaa <- 10
p0mat <- t(array(exp(-0.5*(yvec-ya)^2/sigmaa^2), dim=dim(amat)))
image(xvec, yvec, p0mat, main=bquote(pi(x[0]) %prop% N[y](.(ya),.(sigmaa)^2)),
      xlab=expression(x[0]), ylab=expression(y[0]))
```

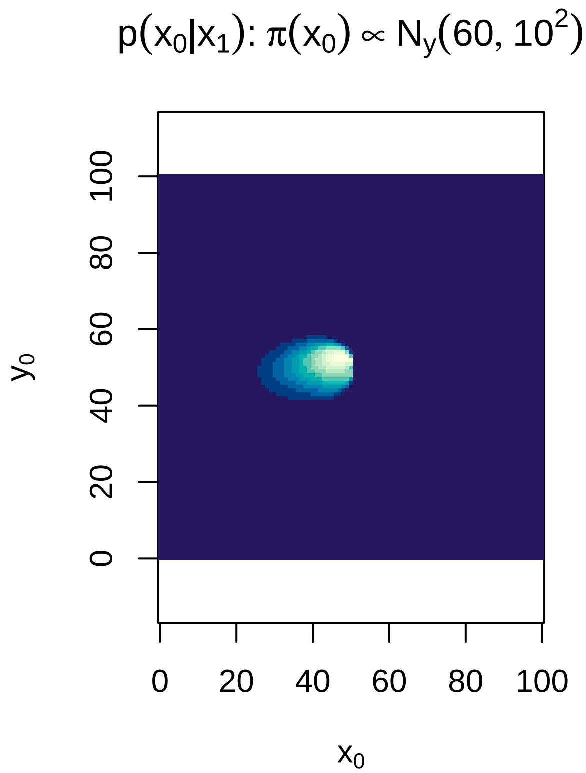


The effect of the prior is shown at right below

```r
par(mfrow=c(1,2))
image(xvec, yvec, qmat, asp=1,
      main=bquote(p(x[0]*"|"*x[1])*":"  ~ pi(x[0]) %prop% 1),
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
image(xvec, yvec, p0mat*qmat, asp=1,
      main=bquote(p(x[0]*"|"*x[1])*":" ~ pi(x[0]) %prop% N[y](.(ya),.(sigmaa)^2)),
      xlab=expression(x[0]), ylab=expression(y[0]), col=colvec)
```



p(x0|x1): π(x0) ∝ 1
y0
100
80
60
40
20
0
0
20
40
60
80
100
x0



p(x0|x1): π(x0) ∝ Ny(60, 10^2)
y0
100
80
60
40
20
0
0
20
40
60
80
100
x0