

Medidas de acción ante roturas de tuberías de agua

Alejandro Vasquez Gavancho¹, Carlos E. Moscol Durand², Thonny J. Lopez del Rio³, Duberly O. Mamani Machaca⁴, Kenjhy J. Bazan Turín⁵, León A. Terrones Caccha⁶ y Docente Jonathan A. Munguia La Cotera⁷

¹Facultad de ciencias, Universidad Nacional de Ingeniería, e-mail: avazquez@uni.pe

²Facultad de ciencias, Universidad Nacional de Ingeniería, e-mail: cmoscold@uni.pe

³Facultad de ciencias, Universidad Nacional de Ingeniería, e-mail: thonny.lopez.d@uni.pe

⁴Facultad de ciencias, Universidad Nacional de Ingeniería, e-mail: duberly.mamani.m@uni.pe

⁵Facultad de ciencias, Universidad Nacional de Ingeniería, e-mail: kbazant@uni.pe

⁶Facultad de ciencias, Universidad Nacional de Ingeniería, e-mail: leon.terrones.c@uni.pe

⁷Facultad de ciencias, Universidad Nacional de Ingeniería, e-mail: jmunguia@uni.pe

Resumen

En los últimos años, científicos e ingenieros han dedicado varios esfuerzos para evitar o detectar un desastre oportunamente a fin de tomar medidas. El objetivo del presente trabajo es proponer un modelo probabilístico que sirva para detectar el desbordamiento de agua debido a roturas o fugas.

Palabras Claves: *Prevención, rotura de tuberías, fugas en tuberías.*

Abstract

In recent years, scientists and engineers have dedicated several efforts to prevent or detect a disaster in time to take action. The objective of the present work is to propose a probabilistic model that serves to detect water overflow due to pipe breaks or pipe leaks.

Keywords: *Prevention, pipe breaks, pipe leaks*

1. INTRODUCCIÓN

Las empresas que gestionan el abastecimiento de agua, como Sedapal (la cual provee de agua a casi 9 millones de peruanos), tienen además el objetivo de operar correctamente las centrales (plantas de agua) con el fin de prestar un adecuado servicio al consumidor ([1] Marsh Perú S.A, 2018). El agua incluso debe tener un determinado caudal y presión, así como cumplir con determinados estándares de calidad. Sin embargo, los sistemas de distribución primarios poseen desperfectos, lo que trae como consecuencia roturas de tubería que impiden el cumplimiento de las condiciones mencionadas (es decir, baja presión y agua posiblemente contaminada). Asimismo, hasta ahora las fugas de agua han causado grandes pérdidas de dinero. Debido al costo tanto social como económico, se ve necesario resolver el problema de las roturas de manera ingeniosa, ya que una sustitución completa de toda la red de tuberías es

inviabile.

Por tal motivo, Sedapal posee sistemas de control que “pueden detectar incrementos fuertes de caudales que pueden ser originados por roturas de tuberías, así como disminuciones de presión de servicio pudiéndose atender de inmediato mediante el telemando de válvulas” ([2] Marsh Risk Consulting, 2018). No obstante, las fugas por rotura de tuberías y encharcamientos persisten, ver Figura 1. Por ello los gestores del agua requieren de modelos que les ayuden en la toma de decisiones, a detectar y prevenir si es posible una rotura de tubería que resulta ser tan perjudicial para la empresa como para los ciudadanos.

El objetivo del modelo será estimar estadísticamente las cotas que delimiten el comportamiento normal del caudal de la tubería (antes de una rotura) en función del tiempo versus un comportamiento anómalo que pudiese provocar una posible fuga de agua, to-

mando como base una muestra histórica de caudal y tiempo. El modelo resultante nos ayudará a tomar una decisión cuando estemos frente a una curva de caudal vs. tiempo real obtenida de realizar suficientes mediciones de caudal en el extremo final de la sección de tubería considerada en riesgo al comparar la evolución del caudal con los datos históricos y el efecto teórico que tendría la rotura. Más específicamente, las cotas formarán regiones dentro del plano caudal vs. tiempo que nos ayudarán a determinar cuantitativamente si la tubería merece una inspección o cualquier acción concreta de prevención. Finalmente, el modelo será puesto en práctica con una serie de valores de caudal generados aleatoriamente y un algoritmo que automatizará la toma de decisiones.

Nuestra aproximación estadística se basa en la aplicación de una distribución normal bivariada para los datos recogidos de caudal en distintos tiempos y está inspirado en el trabajo de [3] Yue (1999) en el análisis de frecuencia de inundaciones a partir de la correlación entre los puntos de caudal más altos y su duración.



Figura 1: Rotura de tubería

2. CONCEPTOS PREVIOS

Probabilidad y estadística

Empezaremos revisando conceptos teóricos de probabilidad y estadística que serán de utilidad en el análisis del problema. Estos conceptos son adaptados de ([4] Ross, 2014; y [5] Moya, 2007)

Definición 1 (Variable aleatoria). Una variable aleatoria \mathbf{X} es una función que asigna un valor a cada posible resultado de un experimento aleatorio de espacio muestral Ω .

$$X : \Omega \rightarrow R_X \quad (1)$$

Existen variables aleatorias cuyo conjunto de posibles valores es finito o infinito numerable, a las cuales llamaremos **variables aleatorias discretas**.

También existen variables aleatorias cuyo rango es un conjunto incontable, llamadas **variables aleatorias continuas**.

Vamos a enfocarnos en esta última pues la variable volumen que vamos a estudiar sigue estas características.

Definición 2. [Función de distribución acumulada] Sea \mathbf{X} una variable aleatoria continua, se define

la función de distribución acumulada, o simplemente función de distribución, como

$$F_X : R_X \rightarrow [0, 1] \\ x \mapsto P(X \leq x) := \int_{-\infty}^x f_X(t) dt, \quad (2)$$

donde f_X es llamada función de densidad.

También se representa la probabilidad de que la variable aleatoria \mathbf{X} esté en un intervalo $[a, b] \subset R_X$ como:

$$P(a \leq X \leq b) := \int_a^b f_X(t) dt \quad (3)$$

Observación 1. La función de distribución se puede ver como una antiderivada de la función de densidad:

$$f_X(x) = \frac{d}{dx} F_X(x) \quad (4)$$

Intuitivamente, la función de distribución de probabilidad $f_X(x)$ nos da la probabilidad de que x esté contenido en un intervalo de longitud infinitesimal alrededor de x .

Definición 3. [Esperanza] Sea \mathbf{X} una variable aleatoria continua con rango R_X y función de densidad f_X , el valor esperado o esperanza matemática de \mathbf{X} se denota por $E[X]$ y se define como

$$E[X] := \int_{\mathbb{R}} t f_X(t) dt. \quad (5)$$

Siempre que la expresión sea absolutamente convergente.

La esperanza matemática de \mathbf{X} se llama también media de la variable aleatoria y se denota por μ .

Definición 4. [Varianza] La varianza de una variable aleatoria \mathbf{X} se denota por $\text{Var}(X)$ o por σ_X^2 (o simplemente σ^2) y se define como

$$\text{Var}(X) := E[(X - \mu)^2], \quad (6)$$

lo cual, se puede expresar como

$$\text{Var}(X) = \int_{\mathbb{R}} (t - \mu)^2 f_X(t) dt. \quad (7)$$

Definición 5. [Desviación estándar] La desviación estándar de una variable aleatoria \mathbf{X} , se denota por σ y se define como la raíz cuadrada de la varianza, es decir:

$$\sigma := \sqrt{\text{Var}(X)} \quad (8)$$

La unidad de σ es la misma que el de la variable aleatoria \mathbf{X} .

Cuando tengamos en contexto más de una variable aleatoria nos referimos a la media, varianza o desviación estándar de X como μ_X , σ_X^2 y σ_X respectivamente, para evitar ambigüedad.

Definición 6. [Función de distribución conjunta] Sean las variables aleatorias continuas X_1 y X_2 , su función de distribución conjunta es

$$F_{X_1, X_2}(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X,Y}(x, y) dy dx \quad (9)$$

Definición 7. [Variables aleatorias independientes] X e Y son variables aleatorias independientes si y solo si

$$F_{X,Y}(x, y) = F_X(x)F_Y(y) \quad (10)$$

para todo x, y real.

Teorema 1. Sean X una variable aleatoria continua, el valor esperado de la función de una variable aleatoria $H(X)$ es

$$E[H(X)] = \int_{-\infty}^{\infty} H(X)f_X(x) \quad (11)$$

Si la expresión es absolutamente convergente

Teorema 2. Sean X e Y dos variables aleatorias, el valor esperado de la función de dos variables aleatorias $H(X, Y)$ es

$$E[H(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(X, Y)f_{X,Y}(x, y) \quad (12)$$

Si la expresión es absolutamente convergente

Teorema 3. Si X e Y son dos variables aleatorias y H y G son funciones de X e Y respectivamente, entonces

$$E[aH(X) + bG(Y)] = aE[H(X)] + bE[G(Y)] \quad (13)$$

Teorema 4. Si X e Y son dos variables aleatorias independientes, entonces

$$1. \quad E[XY] = E[X]E[Y] \quad (14)$$

$$2. \quad \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) \quad (15)$$

La distribución de probabilidad continua más importante en todo el campo de la estadística es, con seguridad, la distribución normal, debido a que en la práctica muchos fenómenos industriales, científicos o de la vida diaria pueden describirse por esta distribución. También mediante esta distribución se obtienen aproximaciones a otras leyes de probabilidad.

Definición 8. [Distribución normal] Una variable aleatoria continua X , se dice que está distribuida normalmente, con media μ y varianza $\sigma^2 > 0$, si su función de densidad de probabilidad está dado por

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (16)$$

La variable aleatoria X que sigue una distribución normal con media μ y varianza σ^2 se representa como

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

Y se puede verificar que:

$$E[X] = \mu \quad (17)$$

$$\text{Var}(X) = \sigma^2 \quad (18)$$

Teorema 5. [Teorema del límite central] Sean X_1, X_2, \dots una secuencia de variables aleatorias independientes e idénticamente distribuidas, cada una con media μ y varianza σ^2 . Entonces se cumple que para $n \rightarrow \infty$:

$$\frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(\mu, \sigma^2) \quad (19)$$

Una consecuencia importante del teorema del límite central es que podemos aproximar la distribución de una muestra aleatoria lo suficientemente grande como una distribución normal. Si hubiera más de un tipo de datos, de modo que estos forman un vector aleatorio, este seguirá una distribución normal multivariada.

Definición 9 (Intervalo de confianza). Sea X_1, X_2, \dots, X_n una muestra aleatoria con media μ , varianza σ^2 y desviación estándar σ . A cada x en el dominio de la distribución normal de la muestra le corresponde un valor u de confianza tal que para un nivel de confianza γ :

$$x - z\frac{\sigma}{\sqrt{n}} \leq u \leq x + z\frac{\sigma}{\sqrt{n}} \quad (20)$$

Donde:

- z es el valor z del nivel de confianza definido como:

$$P(-z \leq X \leq z) = \gamma \quad (21)$$

Donde $X \sim \mathcal{N}(0, 1)$.

- n es el número de observaciones de la muestra

Definición 10 (Esperanza condicional). Sean X e Y variables aleatorias y sea y un número real tal que $P[Y = y] > 0$. El valor esperado condicional a X dado $Y = y$ se denota por $\mu_{X|Y}(y)$ y se define como

$$\begin{aligned} \mu_{X|Y}(y) &= E[X|Y = y] = \int_{R_X} xP(X = x|Y = y) \\ &= \int_{R_X} x p_{X|Y}(x|y) \quad (22) \end{aligned}$$

Definición 11 (Varianza condicional). Sean X e Y variables aleatorias y sea y un número real tal que $P[Y = y] > 0$. La varianza condicional a X dado $Y = y$ se denota por $\sigma_{X|Y}(y)$ y se define como

$$\begin{aligned} \sigma_{X|Y}^2(y) &= E[(X - \mu_{X|Y}(y))^2|Y = y] \\ &= \int_{R_X} (x - \mu_{X|Y}(y))^2 p_{X|Y}(x|y) \quad (23) \end{aligned}$$

Observación 2. Como en el caso de la varianza ordinaria, puede usarse la propiedad

$$\sigma_{X|Y}^2(y) = E[X^2|Y = y] - (E[X|Y = y])^2$$

Las propiedades de la esperanza condicional son completamente similares a las de la esperanza ordinaria.

Teorema 6. Sean X e Y variables aleatorias y sean a y b números reales. Sea $H(X, Y)$ una función de dos variables, entonces:

1.

$$E[aX + b|Y = y] = a E[X|Y = y] + b, \quad \forall y : p_Y(y) > 0 \quad (24)$$

2.

$$E[H(X, Y)|Y = y] = \int_{R_X} H(x, y) p_{X|Y}(x|y) dx \quad (25)$$

Observación 3. A la n -tupla de variables aleatorias (X_1, X_2, \dots, X_n) también se le conoce como vector aleatorio. La función de probabilidad del vector aleatorio es equivalente a la función de distribución conjunta de las n variables aleatorias.

Teorema 7. Si X e Y son variables aleatorias independientes, entonces

$$E[X|Y = y] = E[X], \quad \forall y : p_Y(y) > 0 \quad (26)$$

Definición 12 (Covarianza). Sean X e Y dos variables aleatorias con media μ_X y μ_Y , respectivamente. La covarianza de X e Y , denotado por $\text{Cov}(X, Y)$ o σ_{XY} , se define por

$$\text{Cov}(X, Y) := E[(X - \mu_X)(Y - \mu_Y)] \quad (27)$$

Teorema 8. La covarianza de dos variables aleatorias X e Y con media μ_X y μ_Y , respectivamente, está dado por

$$\text{Cov}(X, Y) = E[XY] - \mu_X \mu_Y \quad (28)$$

Teorema 9. Si X e Y son variables aleatorias independientes, entonces

$$\text{Cov}(X, Y) = 0 \quad (29)$$

Teorema 10. Si X e Y son variables aleatorias con función de probabilidad conjunta y varianzas finitas, entonces

$$\text{Var}(aX \pm bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) \pm 2ab \text{Cov}(X, Y) \quad (30)$$

Definición 13 (Coeficiente de correlación de Pearson). Sean X e Y variables aleatorias con desviación estándar σ_X y σ_Y , respectivamente. El coeficiente de correlación de X e Y , denotado $\rho(X, Y)$, se define por

$$\rho(X, Y) := \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (31)$$

Teorema 11. Para el coeficiente de correlación ρ de las variables aleatorias X e Y se cumplen las siguientes propiedades:

$$1. -1 \leq \rho \leq 1$$

$$2. \rho(X, Y) = \rho(Y, X)$$

$$3. \text{ Si } X \text{ e } Y \text{ son independientes, entonces } \rho = 0$$

$$4. \text{ Si } \rho = \pm 1, \text{ entonces entre } X \text{ e } Y \text{ existe una dependencia funcional lineal}$$

$$5. \text{ Suponga que } Y = aX + b, \text{ donde } a \text{ y } b \text{ son constantes.}$$

$$\blacksquare \text{ Si } a > 0, \text{ entonces } \rho(X, Y) = 1;$$

$$\blacksquare \text{ Si } a < 0, \text{ entonces } \rho(X, Y) = -1$$

$$6. \text{ Si } a, b, c, d \text{ son constantes, } a > 0 \text{ y } b > 0, \text{ entonces;}$$

$$\rho(aX + c, bY + d) = \rho(X, Y)$$

Teorema 12. Sea (X, Y) un vector aleatorio que sigue una distribución normal bivariada, entonces la variable aleatoria condicionada $[X|Y = y]$ sigue una distribución normal con media y varianza conocidos.

$$[X|Y = y] \sim \mathcal{N}\left(\mu_X + \frac{\sigma_X}{\sigma_Y} \rho(y - \mu_Y), (1 - \rho^2) \sigma_X^2\right) \quad (32)$$

Donde $\rho = \rho(X, Y)$.

Física

A continuación, revisaremos algunas leyes físicas que serán de utilidad en el planteamiento del modelo.

Ley de continuidad. La masa de un fluido en movimiento no cambia al fluir. Esto conduce a una relación cuantitativa importante llamada ecuación de continuidad. Considere una porción de un tubo de flujo entre dos secciones transversales estacionarias con áreas A_1 y A_2 . Los valores de la rapidez del fluido en estas secciones son v_1 y v_2 , respectivamente. No fluye fluido a través de los costados del tubo porque la velocidad del fluido es tangente a la pared en todos sus puntos. Consideremos además una densidad constante. Asimismo, se concluye que $dV = A_1 v_1 dt$. De esto, $dm_1 = \rho A_1 v_1 dt$.

$$A_1 v_1 dt = A_2 v_2 dt \quad (33)$$

El producto Av es la tasa de flujo de volumen (caudal).

$$Q = \frac{dV}{dt} = Av \quad (34)$$

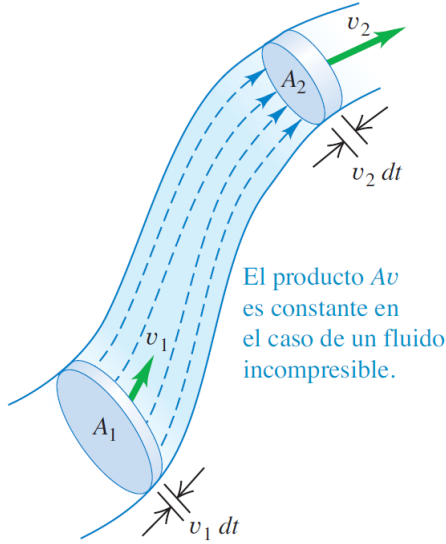


Figura 2: Con densidad constante, el flujo es constante ([6] Young & Freedman, 2012).

1. Ley de continuidad para tuberías ramificadas. Dada las condiciones mencionadas en la ley de continuidad, es posible obtener una generalización para una tubería con n ramificaciones.

$$Q_0 = \sum_{i=1}^n Q_i \quad (35)$$

donde Q_0 es el caudal que entra y Q_i son los caudales en las distintas ramificaciones finales de la tubería ([7] Cabrera, 2007).

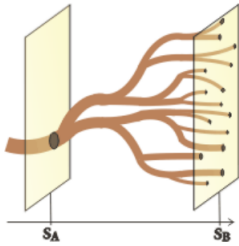


Figura 3: Tuberías ramificadas

3. ANÁLISIS

Modelo físico

Diferentes factores afectan la rotura de tuberías (fugas de agua) ([8] Ávila & González, 2006). Entre los principales se puede mencionar el efecto de las altas presiones pues al aumentar la presión aumenta la velocidad de salida a través del orificio por el que se causa la fuga y con ello el caudal de la misma. Podemos pensar en abstraer el efecto del orificio a un orificio circular, en el que existe una relación entre el caudal de fuga y la presión en la tubería. Entonces para una tubería (en este caso de polietileno de PVC) situada en cualquier

lugar de un sistema de alcantarillado (como se muestra en la figura 1) se tiene la siguiente relación :

$$Q = cP^b \quad (36)$$

Donde:

- Q : caudal,
- c : coeficiente global del cuerpo de descarga,
- P : presión en metros de agua,
- b : exponente en función del diámetro y el espesor de la falla u orificio.

Luego, se calibran los parámetros c y b :

Para el parámetro c , se hace una parametrización en función de la longitud:

$$c = 875L^{1,53} \quad (37)$$

De igual manera, el exponente b en función de la longitud sería:

$$b = 2,53L^{0,18} \quad (38)$$

Y tomando en cuenta el diámetro de la falla, tendríamos la siguiente calibración en b :

$$b = 1,678\left(\frac{L}{d}\right)^{0,03} \quad (39)$$

Aproximándolo sería:

$$b = 1,70 \quad (40)$$

Por lo tanto, se tendría hasta el momento:

$$Q = 875 \times L^{1,53} P^{2,53L^{0,18}} \quad (41)$$

Teniendo en cuenta, nuevamente, el diámetro de la falla:

$$Q = 850d^{-0,7} L^{2,65} P^{1,678\left(\frac{L}{d}\right)^{0,03}} \quad (42)$$

Aproximando el parámetro b se obtiene que la relación entre el caudal y la presión en la salida del orificio, se tiene:

$$Q = 850 \times D^{-0,9} L^{2,70} P^{1,70} \quad (43)$$

Donde :

- Q : caudal (litros/s),
- D : diámetro de la falla (mm),
- L : espesor de la falla (mm),
- P : presión en metros de agua.

Además el caudal también se puede definir como el volumen de agua que atraviesa una superficie en un tiempo determinado.

$$Q = \frac{V}{T} \quad (44)$$

Donde:

- V : volumen (litros),

- T : tiempo (s).

Es complicado encontrar como influye la presión en el caudal con el paso del tiempo ya que la presión depende de muchos factores físicos externos a la tubería. Aún así, de (43) podemos acotar el impacto que tiene en el desbordamiento del agua, con lo que podríamos predecir el instante de tiempo (**tiempo de reacción**) en el cual el caudal pasa a tener un comportamiento anómalo que pueda indicar una fuga y tomar las acciones necesarias.

Finalmente, para realizar las mediciones correspondientes de caudal y tiempo se necesitarán un caudalímetro y un cronómetro, respectivamente. Existe una amplia variedad de instrumentos para calcular el caudal de una tubería que aprovechan distintos principios físicos.

Modelo Matemático

Para empezar, tenemos que definir en términos precisos nuestro objeto de estudio. El modelo se construye a partir de una base estadística de mediciones históricas de caudal en función del tiempo, medidos en un punto determinado de la tubería. La rotura que intentaremos ubicar ocurre entonces en un punto anterior a este que denominaremos punto de rotura. El tramo de tubería estudiado va desde un punto arbitrario anterior al punto de rotura hasta el punto de medición, siempre y cuando no existan ramificaciones en dicho tramo. Además, todas las mediciones actuales que utilicen el modelo se realizarán en un punto posterior al punto de rotura.

De acuerdo a la ecuación de continuidad, el caudal en cualquier sección del tramo estudiado es constante si no existen ramificaciones. Esto no será igual luego de una rotura ya que el orificio dentro de la tubería provocará que parte del volumen de agua transportado sea liberado, lo que provoca una pérdida de caudal. Luego de la rotura, llamaremos caudal de entrada y caudal de salida al valor del caudal antes y después del punto de rotura, denotados por Q y Q_{final} respectivamente. También denominaremos Q_0 al caudal perdido por la rotura. Entonces, tenemos la siguiente ecuación:

$$Q = Q_{\text{final}} + Q_0 \quad (45)$$

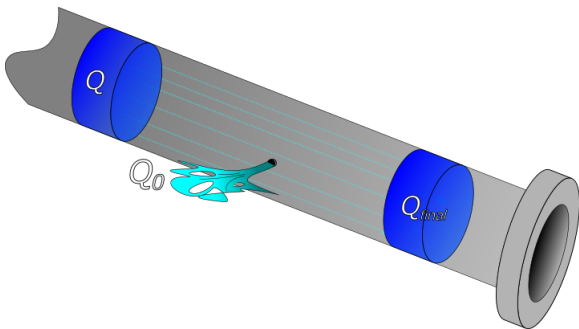


Figura 4: Las roturas generan pérdidas de caudal. El caudal de entrada coincidirá con el caudal regular si se mide en un tramo posterior a la rotura.

Si no hay rotura, entonces $Q_0 = 0$. Podemos considerar entonces que Q es igual al caudal en condiciones normales de la tubería, es decir, antes de que haya una rotura; ya que el que exista o no una rotura no alterará el valor del caudal de entrada. Denominaremos caudal regular al caudal de la tubería antes de la rotura, constante en cualquier sección transversal del tramo según la ecuación de continuidad.

Lo primero que trataremos en la elaboración del modelo será encontrar una estimación estadística de la dependencia entre el caudal regular de la tubería y el tiempo para la base histórica de datos, lo que nos permitirá acotar el caudal. De manera similar tomando en cuenta los resultados experimentales citados en el modelo físico, obtendremos una estimación teórica del efecto de una rotura sobre el caudal histórico con lo que también podremos acotar el caudal final Q_{final} luego de una aparente rotura.

El análisis matemático se construirá a partir de las siguientes suposiciones:

1. La tubería es idéntica en todo el tramo considerado, es decir, tiene el mismo diámetro y el mismo material, con la misma resistencia.
2. La rotura de la tubería será ocasionada únicamente por la presión provocada por el volumen de agua que transita en ella, que estará en función de diferentes parámetros físicos relativos a la tubería (ver modelo físico).
3. Para un registro de datos históricos lo suficientemente grande el caudal, conjuntamente con el tiempo, van a seguir una distribución normal bivariada.

Entonces, definimos la función de probabilidad acumulada conjunta $F_{QT}(q, t)$ como la probabilidad de observar un caudal regular menor o igual a q y un tiempo menor o igual a t . Como asumimos que estos valores siguen una distribución normal bivariada, tendríamos

$$F_{Q,T}(q, t) = \int_{-\infty}^q \int_{-\infty}^t f_{Q,T}(q, t) dt dq,$$

donde:

$$f_{Q,T}(q, t) = \frac{1}{2\pi\sigma_Q\sigma_T\sqrt{1-\rho^2}} e^M, \quad (46)$$

$$M = -\frac{1}{2} \left[\frac{(q - \mu_Q)^2}{\sigma_Q^2} + \frac{(t - \mu_T)^2}{\sigma_T^2} - \frac{2\rho(q - \mu_Q)(t - \mu_T)}{\sigma_Q\sigma_T} \right].$$

Necesitamos conocer como varía el caudal regular en función del tiempo, así que no vamos a considerar la función de distribución conjunta; en lugar de eso analizaremos el caudal condicionado en un instante de tiempo t .

Definimos entonces la función de distribución de probabilidad Q condicionada por T como $f_{Q|T}(q|t)$. Por el Teorema 12, se tiene que esta distribución también es normal

$$[Q|T = t] \sim \mathcal{N}\left(\mu_Q + \frac{\sigma_Q}{\sigma_T}\rho(t - \mu_T), (1 - \rho^2)\sigma_Q^2\right). \quad (47)$$

Si bien el conjunto de posibles valores para el caudal en un tiempo fijo no está acotado y este podría tomar cualquier valor, vamos a considerar un intervalo de confianza alrededor de la media de la distribución de Q con un nivel de confianza del 95 %. De este modo el intervalo de valores para Q está acotado y podemos dar una aproximación útil de como cambia el caudal regular con el paso del tiempo. Así, nuestra aproximación estará acotada de la siguiente forma:

$$\boxed{\mu_{Q|T}(t) - r \leq Q(t) \leq \mu_{Q|T}(t) + r}, \quad (48)$$

donde:

- r es el margen de error,

$$r = z \frac{\sigma_{Q|T}}{\sqrt{n}} \quad (49)$$

- z es el valor del nivel de confianza
- n es el número de observaciones consideradas.

Observación 4. Según (47) solo la media de la distribución normal bivariada depende del tiempo; mientras que la varianza es siempre constante en el tiempo. Por eso se omite el parámetro t en $\sigma_{Q|T}(t) = \sigma_{Q|T}$.

Definimos S y S^* como las curvas que acotan de manera superior e inferior el caudal regular Q , respectivamente. Es decir:

$$S^* : \frac{\sigma_Q}{\sigma_T} \rho(t - \mu_T) + \mu_Q - r \quad (50)$$

$$S : \frac{\sigma_Q}{\sigma_T} \rho(t - \mu_T) + \mu_Q + r \quad (51)$$

Observamos claramente que las curvas S y S^* representan dos rectas de pendiente igual a $\frac{\sigma_Q}{\sigma_T} \rho$. Estas curvas encierran una región en el plano caudal vs. tiempo, que denominaremos **región regular** de la función caudal. Se espera que, en condiciones normales, la función $Q(t)$ esté acotada en dicha región. También llamaremos a la recta S^* como la **curva de alerta**

Ahora que el caudal regular Q ya está acotado, buscamos acotar el caudal luego de una rotura Q_{final} . De la ecuación (45) tenemos que:

$$Q_0 = Q - Q_{\text{final}} \quad (52)$$

Donde Q_0 es la pérdida de caudal debido a la rotura. De la ecuación (43) sabemos que el caudal de fuga (en m^3/s) para un orificio circular está determinado en función de la presión en la sección transversal correspondiente a la rotura

$$Q_0 = 0,85 \times D^{-0,9} L^{2,70} P^{1,70} \quad (53)$$

Donde D y L son el diámetro y el espesor del orificio en milímetros respectivamente.

Es complicado saber como varía dicha presión en función del tiempo, pues en ella influyen muchos factores físicos externos. Aún así, podemos encontrar una

cota mínima pues sabemos que la presión siempre será mayor que la presión atmosférica P_a , que equivale aproximadamente a 10,33 metros de agua. Por lo tanto tenemos la desigualdad:

$$\begin{aligned} Q_0 &\geq 0,85 \times D^{-0,9} L^{2,70} P_a^{1,70} \\ \therefore Q_0 &\geq 45,06 \times D^{-0,9} L^{2,70} \end{aligned} \quad (54)$$

De la ecuación (52) tenemos:

$$\begin{aligned} Q - Q_{\text{final}} &\geq 45,06 \times D^{-0,9} L^{2,70} \\ \therefore Q_{\text{final}}(t) &\leq Q(t) - 45,06 \times D^{-0,9} L^{2,70} \end{aligned} \quad (55)$$

Además, de la desigualdad en (48) tenemos que

$$Q(t) \leq \mu_{Q|T}(t) + r \quad (56)$$

Por lo tanto, de las desigualdades (55) y (56) obtenemos la desigualdad:

$$Q_{\text{final}}(t) \leq \mu_{Q|T}(t) + r - 45,06 \times D^{-0,9} L^{2,70} \quad (57)$$

Con lo que llegamos a una cota superior de Q_{final} . Por simplicidad, suponemos que D y L son constantes en el tiempo.

Para que Q_{final} esté acotada inferiormente, dado que es complicado hallar una cota inferior no trivial, nos basaremos en un nivel de riesgo aceptable que considere quien utiliza el modelo.

Definimos el coeficiente k como el cociente mínimo que aceptaremos entre el caudal de salida y el caudal de entrada. Es decir, k es un número entre 0 y 1 tal que:

$$\frac{Q_{\text{final}}}{Q} \geq k \quad (58)$$

Observación 5. El número k no es un valor que necesariamente satisfaga la desigualdad anterior para cualquier instante de tiempo. Usaremos esta distinción entre valores de t que satisfacen o no esta característica para tomar la decisión entre tomar acciones para evitar un perjuicio de fuga mayor o no. Esto se detallará más adelante.

También podemos encontrar una relación similar para la pérdida de caudal Q_0 . A partir de las expresiones (52) y (58), obtenemos:

$$\frac{Q_0}{Q} \leq 1 - k \quad (59)$$

Entonces, definimos el **coeficiente de desbordamiento** d :

$$d := 1 - k \quad (60)$$

El número d representa el cociente máximo que aceptaremos entre la pérdida de caudal y el caudal de entrada.

De las desigualdades en (48) y (58), obtenemos de manera directa la desigualdad:

$$Q_{\text{final}}(t) \geq k(\mu_{Q|T}(t) - r) \quad (61)$$

donde $k = 1 - d$

Sea R la curva de riesgo que representa a la cota superior de Q_{final} y sea R^* la **curva de riesgo límite** que representa a la función $(1-d)(\mu_{Q|T}(t) - r)$. De forma explícita escribimos:

$$R^* : (1-d) \left(\frac{\sigma_Q}{\sigma_T} \rho (t - \mu_T) + \mu_Q - r \right) \quad (62)$$

$$R : \frac{\sigma_Q}{\sigma_T} \rho (t - \mu_T) + \mu_Q + r - 45,06 \times D^{-0,9} L^{2,70} \quad (63)$$

Observamos que R y R^* representan rectas de pendiente $\frac{\sigma_Q}{\sigma_T} \rho$ y $k \frac{\sigma_Q}{\sigma_T} \rho$, respectivamente. Estas curvas encierran una región en el plano caudal vs. tiempo, que denominaremos **región de riesgo** de la función caudal.

La región de riesgo se interpreta como el espacio en el que debe estar acotada la función $Q_{\text{final}}(t)$ antes de que la rotura de la tubería provoque un desbordamiento o pérdida de caudal mayor a d veces el caudal de entrada $Q(t)$ o, en otras palabras, a un porcentaje d de Q .

Dado que no es posible medir de forma continua el caudal, en la actualidad solo será posible calcular un número finito de puntos en el gráfico caudal vs. tiempo. Para calcular una aproximación de los puntos de intersección de $Q(t)$ con las rectas utilizaremos una interpolación lineal.

4. MODELAMIENTO (Algoritmo)

Modelo computacional

Para crear dicho modelamiento será útil el lenguaje de programación estadístico R.

En esta presentación final se utilizarán datos pseudo aleatorios del caudal de agua *en* (m³/h) que pasa por un determinado tramo arbitrario de tubería y un tiempo transcurrido T (en horas), que nos permiten co-tejar nuestro modelo planteado. Sin embargo, en una futura continuación del proyecto utilizaremos una base de datos históricos .

1. Generamos datos pseudoaleatorios que simulan las muestras históricas de las variables **caudal** Q (con distribución normal: media μ y varianza σ^2) y **tiempo** T que pasa por un tramo arbitrario de una tubería.

```
1 set.seed(123) # Produce la misma muestra una y otra vez.
2 caudal = rnorm(6, mean=10, sd=2) # Genera valores con media 10 y desviación estándar 2
3 tiempo = 1:6 # Genera valores
```

Listing 1: Datos Históricos

2. Creamos un DataFrame con las variables **caudal** Q y **tiempo** T .

```
1 dato = data.frame(tiempo, caudal) #Se guardan las muestras de las variables Tiempo y Cantidad
```

Listing 2: DataFrame

3. Se define las funciones para hallar la **varianza** σ^2 y **desviación estándar** σ de las variables caudal y tiempo de forma general.

```
1 variance = function (x) sum((x - mean(x))^2)/(length(x))
2 desviacion = function (x) sqrt(sum((x - mean(x))^2)/(length(x)))
```

Listing 3: Funciones de varianza σ^2 y desviación estándar σ

4. Definimos la **media** μ y las medidas de dispersión: **varianza** σ^2 , **desviación estándar** σ y **coeficiente de correlación** ρ de las variables caudal y tiempo.

```
1 #Distribucion normal
2 n=length(dato$tiempo) #tamano de la muestra
3 med_t=mean(dato$tiempo) #Media del tiempo
4 med_q=mean(dato$caudal) #Media de la caudal
5 var_t=variance(dato$tiempo) #Varianza del tiempo
6 var_q=variance(dato$caudal) #Varianza de la caudal
7 des_t=desviacion(dato$tiempo) #Desviacion del tiempo
8 des_q=desviacion(dato$caudal) #Desviacion de la caudal
9 coef_p=cor(dato$tiempo, dato$caudal) # Coeficiente de correlacion
```

Listing 4: Media y medidas de dispersión de una distribución normal

5. Definimos el **nivel de confianza** y Z .

```
1 nv = 0.95 #Nivel de confianza
2 z = qnorm((nv+1)/2, 0, 1)
```

Listing 5: Nivel de confianza y Z

6. Definimos la **media** (med_bi), **varianza** (var_bi), **desviación estándar** (des_bi) y **margen de error** (r_bi) de la función de distribución normal bivarida Q condicionada por T ,

```
1 t = 2 #lee el tiempo a calcular
2 med_bi = med_q + (des_q/des_t)*coef_p*(t - med_t) #Media bivarida
3 var_bi = (1-(coef_p^2))*var_q #Varianza bivarida
4 des_bi = sqrt((1-(coef_p^2))*var_q) # Desviacion bivarida
5 r_bi = (z*des_bi)/sqrt(n) #Margen de error
```

Listing 6: Función de distribución normal bivarida Q condicionada por T

7. Definimos los límites de la región regular de nuestra distribución normal bivarida condicionada definida en el ítem anterior.


```

1 #Region Regular de la distyribucion
  normal bivariada caudal condicionada
  por el tiempo
2 l_inf_bi = med_bi-r_bi #Limite inferior
  del Intervalo de confianza
3 l_sup_bi = med_bi+r_bi #Limite superior
  del intervalo de confianza
4

```

Listing 7: Región regular

8. Se genera un valor aleatorio con la distribución normal bivariada condicionada en el instante x con respecto al intervalo de confianza.

```

1 set.seed(Sys.time()) #Para que la grafica
  generada sea completamente aleatoria
2
3 mu_biv_dist <- function(x)
4 {
5   return(med_q+(des_q/des_t)*coef_p*(x-
     med_t))
6 }
7
8 #Genera un valor aleatorio con la
  distribucion normal del instante x
9 biv_dist <- function(x)
10 {
11   mu = mu_biv_dist(x)
12   return(rnorm(1, mu, des_bi))
13 }
14 green_line <- function(x)
15 {
16   return (mu_biv_dist(x) - rad_bi)
17 }
18
19 red_line <- function(x)
20 {
21   return ((1-d)*(mu_biv_dist(x) - rad_bi)
22   )
23 }
24 start = 0 # Inicio del dominio de la
  grafica a generar
25 end = 10 # Final del dominio de la
  grafica a generar
26 step = 0.1
27 D = 0.05 # Diametro de la falla
28 L = 0.05 # Espesor de la Falla
29 Pa = 1/0.0967838 #Presion atmosferica
30 P = 0.85*D^(-0.9)*L^(2.70)*Pa^(1.70) #
  Minima presion en la rotura
31
32 d = 0.25 #Coeficiente de desbordamiento
33
34 times = seq(start, end, step)
35 plot_values = sapply(times, biv_dist)
36
37 contador = 0
38 in_region = plot_values[1] > green_line(
  step*(t-1))
39

```

Listing 8: Distribución normal en el instante x

9. Definimos las funciones para hallar los tiempo de alerta y el tiempo crítico.

```

1 #Los tiempos de alerta y critico se
  aproximan con una interpolacion
  lineal
2 hallar_tiempo_alerta <- function(x)
3 {
4   m1 = (plot_values[x] - plot_values[x
     -1])/step

```

```

5   b1 = plot_values[x-1]*(x-1) - plot_
     values[x]*(x-2)
6   m2 = (des_q/des_t)*coef_p
7   b2 = med_q - rad_bi - (des_q/des_t)*
     coef_p*med_t
8
9   return ((b1-b2)/(m2-m1))
10 }
11
12 hallar_tiempo_critico <- function(x)
13 {
14   m1 = (plot_values[x] - plot_values[x
     -1])/step
15   b1 = plot_values[x-1]*(x-1) - plot_
     values[x]*(x-2)
16   m2 = (1-d)*(des_q/des_t)*coef_p
17   b2 = (1-d)*(med_q - rad_bi - (des_q/des
     _t)*coef_p*med_t)
18
19   return ((b1-b2)/(m2-m1))
20 }
21
22 lecturas_criticas = c()
23
24 for (t in seq(2, length(times)))
25 {
26   if ((plot_values[t] > green_line(step*(
     t-1))) != in_region)
27   {
28     if (in_region)
29     {
30       temp = hallar_tiempo_alerta(t)
31       cat("Alerta", temp, "horas\n")
32       append(lecturas_criticas, temp)
33       contador = contador + 1
34     }
35     in_region = !(in_region)
36   }
37   if (plot_values[t] <= red_line(step*(t
     -1)))
38   {
39     tiempo_critico = hallar_tiempo_
     critico(t)
40     cat("\nLecturas criticas:", contador,
41     "\nComportamiento anomalo y
     posible fuga, tiempo: ", tiempo_
     critico,
42     "horas.\nSe debe revisar el
     estado de la tuberia.\n")
43     break
44   }
45 }
46

```

Listing 9: Funciones para hallar el tiempo de alerta y el tiempo crítico

10. Generamos las gráficas de las simulaciones realizadas.

```

1 #MOSTRAMOS LAS GRAFICAS
2 pdf(file="1. sin interpolar.pdf")
3
4 plot(times, plot_values, xlab="Tiempo (
  horas)", ylab="Caudal (m3/s)")
5 lines(c(start, end), c(mu_biv_dist(start)
  - rad_bi, mu_biv_dist(end) - rad_bi)
  , col="green")
6 lines(c(start, end), c(mu_biv_dist(start)
  + rad_bi, mu_biv_dist(end) + rad_bi)
  , col="green")
7 lines(c(start, end), c(mu_biv_dist(start)
  + rad_bi - P, mu_biv_dist(end) + rad
  _bi - P), col="red")
8 lines(c(start,end), c((1-d)*(mu_biv_dist(
  start) - rad_bi), (1-d)*(mu_biv_dist(

```

```

end) - rad_bi)), col="red")
9
10 legend(x = "bottomright", legend = c("
    Region regular", "Region de riesgo"),
    fill = c("green", "red"), title = "
    Datos no interpolados")
11
12 dev.off()
13
14 pdf(file="2. interpolado.pdf")
15
16 plot(times, plot_values, xlab="Tiempo (
    horas)", ylab="Caudal (m3/s)", type="
    l")
17 lines(c(start, end), c(mu_biv_dist(start)
    - rad_bi, mu_biv_dist(end) - rad_bi)
    , col="green")
18 lines(c(start, end), c(mu_biv_dist(start)
    + rad_bi, mu_biv_dist(end) + rad_bi)
    , col="green")
19 lines(c(start, end), c(mu_biv_dist(start)
    + rad_bi - P, mu_biv_dist(end) + rad
    _bi - P), col="red")
20 lines(c(start, end), c((1-d)*(mu_biv_dist(
    start) - rad_bi), (1-d)*(mu_biv_dist(
    end) - rad_bi)), col="red")
21
22 legend(x = "bottomright", legend = c("
    Region regular", "Region de riesgo"),
    fill = c("green", "red"), title = "
    Datos interpolados")
23
24 dev.off()
25

```

Listing 10: Gráficas de las simulaciones realizadas

5. RESULTADOS Y DISCUSIONES

- La mayoría de los programas de computadora utilizan algoritmos deterministas para generar números aleatorios (que es la razón por la cual los números que generan no son verdaderamente aleatorios, sino pseudoaleatorios, lo cual es suficientemente bueno para la mayoría de los propósitos). R no es diferente, y puede pensar en los números aleatorios que genera como parte de una cadena muy larga de números "aleatorios" que, cuando se invoca, simplemente comienza en algún punto y muestra números pseudoaleatorios para usted. Al usar `set.seed()`, básicamente le está dando al programa un punto de partida en lugar de dejar que elija el suyo. Es por eso que cualquier usuario que ejecute el mismo número de semilla obtendrá los mismos resultados.

- Analizamos datos pseudoaleatorios dado que no contamos con los datos históricos, pues dicha información es confidencial para las empresas que gestionan el abastecimiento de agua.

Para nuestro cálculo, supondremos que el caudal observado tiene una media de 10 y desviación estándar de 2, en metros cúbicos por segundo. Entonces simularemos los datos normales para el caudal con valores aleatorios de una distribución normal con $\mu = 10$ y $\sigma = 2$.

- El marco de datos de la variables caudal y tiempo es el siguiente:

```

1 > tiempo    caudal
2 1          1  8.879049
3 2          2  9.539645
4 3          3 13.117417
5 4          4 10.141017
6 5          5 10.258575
7 6          6 13.430130
8

```

Listing 11: DataFrame

- Obtenemos la media, varianza, desviación estándar, y coeficiente de correlación para cada variable.

```

1 cuadro_tendencia = data.frame(
2   "Variables" = c("Tiempo(T)", "Caudal(Q)"),
3   "Media" = c(med_t, med_q),
4   "Desviacion_estandar" = c(des_t, des_q),
5   "Varianza" = c(var_t, var_q),
6   "Coeficiente_de_correlacion_(p)" = c(
7     coef_p
8   )
9   cuadro_tendencia
10
11 > Variables      Media      Des_estandar
12 1 Tiempo(T)    3.50000      1.707825
13 2 Caudal(Q)   10.89431      1.743588
14
15 Varianza      Coef_de_corr_p
16 2.916667      0.6138821
17 3.040098      0.6138821
18

```

Listing 12: Media μ y medidas de dispersión

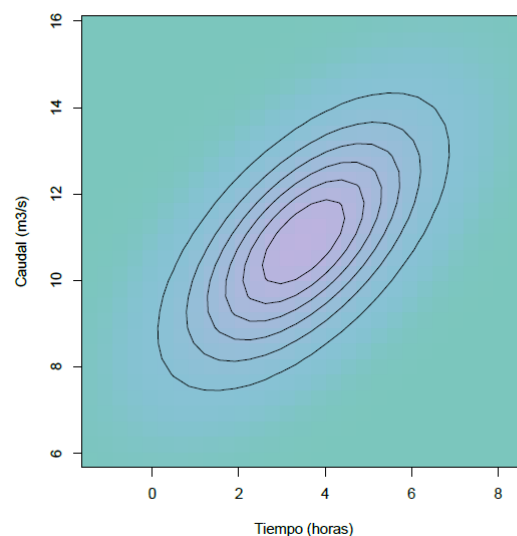


Figura 5: Regiones de confianza de la distribución conjunta de los datos históricos

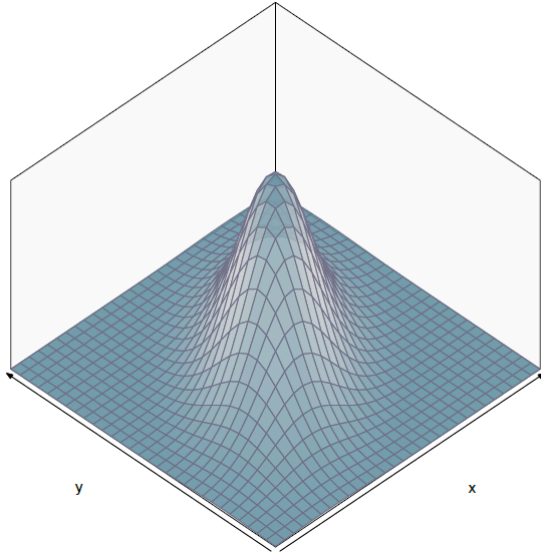


Figura 6: Gráfica de la distribución conjunta de los datos históricos

- Obtenemos el nivel de confianza y el valor de Z:

```
1 cuadro_confianza = data.frame(
2   "Nivel_de_confianza" = c(nv),
3   "Valor_Z" = c(z)
4 )
5 cuadro_confianza
6
7 > Nivel_de_confianza  Valor_Z
8                      0.95  1.959964
9
```

Listing 13: Nivel de confianza y Z

- Obtenemos los parametros de nuestra distribución normal bivariada condicionada.

```
1 #Cuadro condicionado
2 #Esta tabla es referencial para calcular
3 #automaticamente cualquier distribucion
4 #normal bivariada condicionada
5 #correspondiente a cualquier tiempo.
6 cuadro_condicionado = data.frame(
7   "Q|T" = c(t),
8   "Media" = c(med_bi), \textbf{\textit{media \mu}} ,
9   "Desviacion_estandar" = c(des_bi),
10  "Varianza" = c(var_bi),
11  "M._Error" = c(rad_bi),
12  "L_Inferior_IC" = c(l_inf_bi),
13  "L_Superior_IC" = c(l_sup_bi)
14 )
15 cuadro_condicionado
16
17 > Q|T  Media Desviacion_estandar Varianza
18      2  9.9542                1.376384  1.894434
19
20 > M._Error L_Inferior_IC L_Superior_IC
21      1.101317      8.852883      11.05552
22
```

Listing 14: Función de distribución normal bivariada condicionada

- Verificamos las alertas a través de múltiples simulaciones, discutiremos dos de ellas una donde se generen alertas criticas y otras donde no.

Para generar dichas alertas es necesario tomar algunos valores contantes definidos en el item 8 del modelo computacional:

- $D = 0.05$ m, diámetro de la falla.
- $L = 0.05$ m, espesor de la falla.
- $P_a = 1/0.0967838$ mH₂O, presión atmosférica

El algoritmo podría indicar un mayor nivel de riesgo o no según los datos actuales medidos, los cuales modelaremos de modo que sigan la tendencia del caudal regular. Veamos los siguientes escenarios:

Con una semilla cuyo valor es 24, ejecutamos previamente en R la línea `set.seed(24)` y obtenemos la siguientes alertas:

```
1 #Los tiempos de alerta y critico se
   aproximan con una interpolacion
   lineal
2 > "Alerta" 0.7873733 horas
3   "Alerta" 0.9608287 horas
4   "Alerta" 1.205689 horas
5   "Alerta" 2.117655 horas
6   "Alerta" 2.882638 horas
7   "Alerta" 3.484174 horas
8
9 > Lecturas criticas: 6
10 Comportamiento anomalo y posible fuga,
   tiempo: 3.544466 horas.
11 Se debe revisar el estado de la tuberia
12 .
```

Listing 15: Alertas criticas

Con una semilla cuyo valor es 124, ejecutamos `set.seed(124)` y obtenemos la siguientes alertas:

```
1 #Los tiempos de alerta y critico se
   aproximan con una interpolacion
   lineal
2
3 "Alerta" 1.064202 horas
4 "Alerta" 1.660552 horas
5 "Alerta" 2.853985 horas
6 "Alerta" 3.047847 horas
7 "Alerta" 3.792012 horas
8 "Alerta" 4.580848 horas
9 "Alerta" 5.192549 horas
10 "Alerta" 5.696698 horas
11 "Alerta" 6.188915 horas
12 "Alerta" 7.079984 horas
13 "Alerta" 8.31458 horas
14 "Alerta" 8.966141 horas
15 "Alerta" 9.764117 horas
16
```

Listing 16: Alertas no criticas

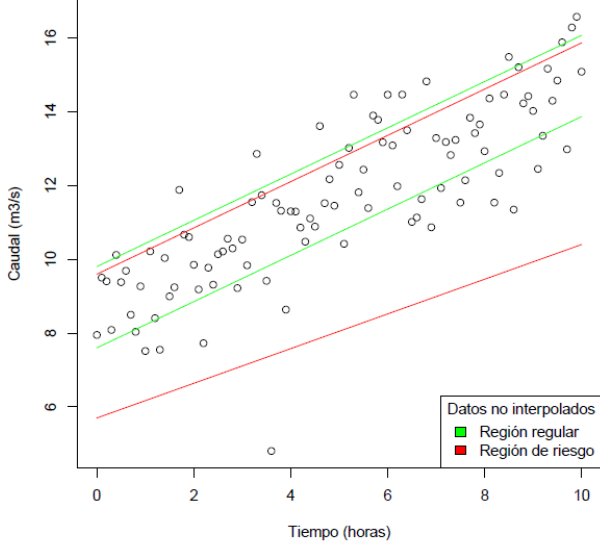


Figura 7: Gráfica (semilla = 24)

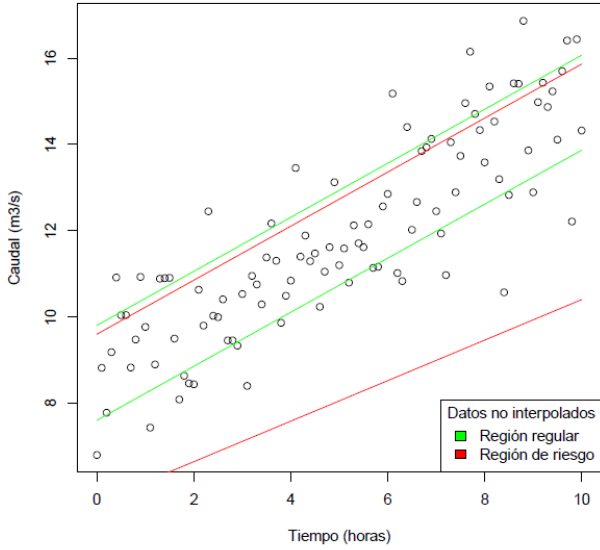


Figura 8: Gráfica (semilla = 124)

- Se puede observar claramente, por la continuidad teórica de la función caudal, que la función real va a intersectar a la recta S^* en más de una ocasión, como nos anticipa el programa cuando se ejecuta el código. También, la función caudal para una semilla igual a 24 va a intersectar una única vez a la recta R^* .
- Como no es posible conocer en qué instantes de tiempo ocurren dichas intersecciones, es decir, en qué momentos se activan las alarmas, se debe aproximar realizando alguna interpolación entre

dos puntos contiguos entre los cuales sabemos que habrá una intersección. Para nuestro modelo, aproximaremos los tiempos de las alarmas con una interpolación lineal de las funciones obtenidas previamente. Así, obtendremos funciones con una apariencia más natural tomando en consideración la naturaleza continua del caudal.

- Las funciones que encuentra la intersección de las rectas de alerta y de riesgo con la interpolación del caudal en puntos contiguos se encuentra en el ítem 9 del modelo computacional. La derivación de estas fórmulas es muy natural. Para ello, podemos parametrizar los puntos discontinuos de la gráfica en función de un número natural t que nos indicará en qué punto (de izquierda a derecha) nos encontramos. Sea (x, y) un punto de la función se puede comprobar que:

$$x(t) = \text{step} * (t - 1) \quad (64)$$

$$y(t) = \text{plot_values}[t] \quad (65)$$

Y a partir de ello se construye la ecuación de la recta entre los puntos contiguos. Sean m_1 y m_2 las pendientes de la recta contigua entre los dos puntos y la recta con la que queremos intersectar y sean b_1 y b_2 los interceptos con el eje de las ordenadas de las mismas rectas respectivamente, con las igualdades de (64), (65), (50) y (62) obtenemos las mismas igualdades del ítem 9 del modelo computacional. A continuación solo queda calcular la intersección de dos rectas dadas sus pendientes y sus ordenadas, por lo cual se redujo a un problema de álgebra muy elemental.

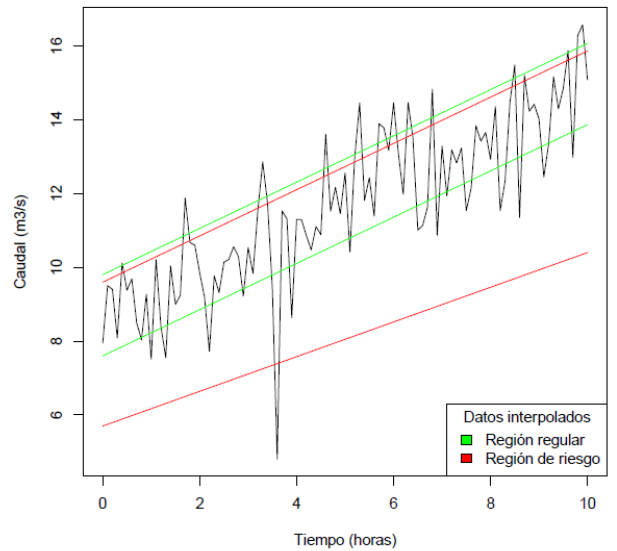


Figura 9: Gráfica con interpolación (semilla = 24)

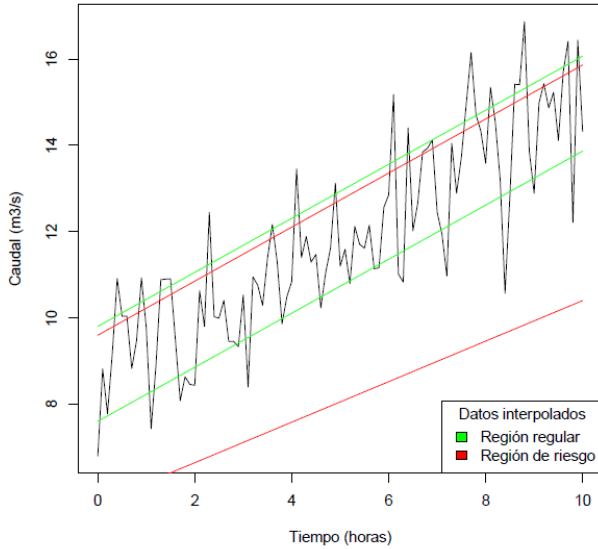


Figura 10: Gráfica con interpolación (semilla = 124)

- Finalmente, los tiempos de alarma que se observan en las gráficas interpoladas coinciden con los tiempo obtenidos en los bloques de código 15 y 16.

6. CONCLUSIONES

- Las regiones regular y de riesgo están acotadas por dos rectas cada una, de las cuales tres de ellas son paralelas entre sí. De sus pendientes se observa que nunca se intersecan en el dominio de la función caudal.
- Al algoritmo solo le importará conocer las ecuaciones de las rectas de alerta y de riesgo límite mientras que las otras no serán evaluadas pues los aumentos de caudal se asume que ocurren de manera natural. Estas rectas se grafican como referencia.

Referencias

- [1] Marsh Perú S.A Corredores de Seguros. (2018). Informe de Suscripción de Riesgos. 05/07/2020, de Marsh Perú S.A. pp. 52-53. Recuperado de http://www.sedapal.com.pe/c/document_library/get_file?uuid=c9cc9dd0-8807-46f2-a46f-b3ee7a5fc674&groupId=10154.
- [2] <https://www.marsh.com/us/services/marsh-risk-consulting.html>
- [3] Yue, S. (1999). Applying Bivariate Normal Distribution to Flood Frequency Analysis. *Water International*, 24(3), 248-254.

- Sin embargo, con respecto al punto anterior, se debe recalcar que la cota superior de la región de riesgo R es una estimación a partir de los parámetros físicos tomados en consideración. Un estudio más cuidadoso y preciso de las condiciones físicas del problema pueden mejorar esta aproximación.
- El modelo propuesto puede funcionar mejor cuando los datos históricos sean más recientes. El caudal no se alejará mucho de su valor promedio y se ajustará mejor al modelo.
- El modelo es netamente estadístico, por lo que pueden detectarse falsas alarmas si la reducción de caudal ocurre de manera natural.

7. SUGERENCIAS

- Los datos que se van a estudiar no siempre se ajustan a una distribución normal y eso puede generar un margen de error en el modelo final. Existen métodos estadísticos para ajustar datos a una distribución normal (normalizar datos) y uno de ellos es la transformación Box-Cox.

Definición 14. Sean y_i un valor de una muestra estadística, entonces $y_i^{(\lambda)}$ es el resultado de aplicarle la transformación Box-Cox ([9] Box & Cox, 1964), definida como:

$$y_i^{(\lambda)} = \begin{cases} \frac{y_i^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln y_i, & \lambda = 0 \end{cases} \quad (66)$$

donde λ es el parámetro de la transformación.

Las muestras transformadas normalizadas $y_i^{(\lambda)}$ se ajustan mejor a la distribución normal y pueden garantizar resultados más eficaces.

- Por último, esperamos que este modelo sea puesto a prueba con datos reales que podamos obtener con las mediciones de caudal obtenidas por hora, que como se mencionó anteriormente, dicha información suele ser confidencial para los sistemas de distribución de agua potable.

- [4] Ross, S. (2014). *A First Course in Probability*. Harlow, Essex: Pearson.
- [5] Moya, R., & Saravia, G. (2007). *Probabilidad e inferencia estadística*. Perú.
- [6] Young, H., & Freedman, R. (2012). *Física Universitaria*, Volumen I, México, Pearson.
- [7] Cabrera, R. (2007). *Hidrodinámica - Caudal en las ramificaciones*. Recuperado de https://ricuti.com.ar/no_me_salene/hidrodinamica/FT_capilares.html
- [8] Ávila, H., & González, C. (2006, noviembre). Determinación de parámetros de fuga para fallas longitudinales, en conexiones domiciliarias y en uniones de tuberías en PVC.

Revista de Ingeniería Universidad de los Andes, núm 24, pp. 15-22. Recuperado de [https://www.researchgate.net/publication/237023400-Determinacion-de-parametros-de-fuga-para-fallas-longitudinales-en-](https://www.researchgate.net/publication/237023400-Determinacion-de-parametros-de-fuga-para-fallas-longitudinales-en-conexiones-domiciliarias-y-en-uniones-de-tuberias-en-PVC)

[conexiones_domiciliarias_y_en_uniones_de_tuberias_en_PVC](https://www.researchgate.net/publication/237023400-Determinacion-de-parametros-de-fuga-para-fallas-longitudinales-en-conexiones-domiciliarias-y-en-uniones-de-tuberias-en-PVC)

- [9] Box, G.E.P., & D.R. Cox. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society*, 26(2), 211-252.