

MODELITZACIÓ I PREDICCIÓ TEMPERATURA I PRECIPITACIONS A ITÀLIA

MARIO LEFEBVRE *

RESUM. Estudiem la temperatura mitjana mensual a Itàlia durant el període 1991-2015. L'augment o la disminució de la temperatura mitjana respecte a l'any anterior es modela com una cadena de Markov de temps discret amb quatre estats possibles. De la mateixa manera, es proposa una cadena de Markov com a model per a les variacions de la quantitat mensual de precipitació. A partir d'aquests models, és possible predir si és probable que la temperatura i la quantitat de precipitació variïn significativament a llarg termini.

1. Introducció

Els processos estocàstics, i en particular les cadenes de Markov, han estat utilitzats per molts autors per modelar diversos fenòmens. Per exemple, Avilés et al. (2016) van predir esdeveniments de sequera basant-se en cadenes de Markov. Drton et al. (2003) van proposar un model de cadena de Markov de l'activitat tornàdica. Matis et al. (1989) van utilitzar aquests processos per predir els rendiments del cotó. Tagliaferri et al. (2016) van generar sèries temporals artificials de velocitat del vent amb l'ajuda de cadenes de Markov imbricades. En hidrologia i meteorologia, els processos de Markov són molt útils per predir els cabals dels rius i les precipitacions (vegeu, per exemple, Caskey Jr. 1963; Lefebvre i Guibault 2008).

El nostre objectiu en aquest article és, en primer lloc, trobar un model estocàstic per a les variacions de la temperatura i la precipitació mitjanes mensuals en el cas d'Itàlia, i després utilitzar aquest model per predir la seva evolució a llarg termini. Es proposarà una cadena de Markov de temps discret amb quatre estats possibles com a model per a la temperatura i també per a la precipitació.

El Banc Mundial ha creat el Portal de Coneixement sobre el Canvi Climàtic. Al lloc web <http://sdwebx.worldbank.org/climateportal>, es poden trobar les temperatures i precipitacions mitjanes mensuals històriques per a tots els països del món. Les dades estan disponibles per al període 1901-2015. Utilitzarem les dades dels anys 1991-2015 per estimar les diverses probabilitats de transició de les cadenes de Markov. A continuació, es calcularan les probabilitats límit de les cadenes de Markov per intentar predir el comportament a llarg termini de les variables d'interès. El conjunt de dades també es dividirà en dues parts per comprovar si hi ha hagut canvis significatius en les variacions de temperatura i precipitació durant el període considerat. A la següent secció, presentarem el model que proposem i, a continuació, el model s'implementarà per a la temperatura i la precipitació a les seccions 3 i 4, respectivament.

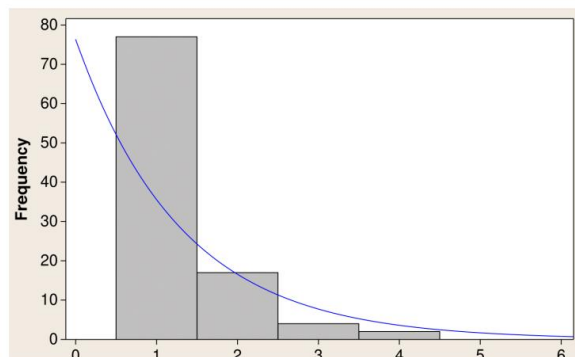


FIGURA 1. Histograma d'una distribució geomètrica.

2. Cadenes de Markov en temps discret

Siguin X_0, X_1, \dots variables aleatòries. Suposem que els valors possibles de les variables aleatòries són els nombres enters $S := \{0, 1, \dots\}$ (on el símbol ":= " significa per definició). El procés estocàstic $\{X_n, n = 0, 1, \dots\}$ és una cadena de Markov (de temps discret) si

$$P[X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0] = P[X_{n+1} = j \mid X_n = i] \quad (1)$$

per a tots els estats $i_0, \dots, i_{n-1}, i, j$ en S , i per a qualsevol $n \in \{0, 1, \dots\}$. És a dir, la probabilitat condicional que el procés estigui en l'estat j en el temps $n+1$, donada tota la història del procés, ja que l'estat inicial i_0 només depèn del valor més recent de l'estat.

Diem que el futur ($n+1$), donat el present (n) i el passat ($n-1, \dots, 0$), només depèn del present. En general, també suposem que

$$p_{ij} := P[X_{n+1} = j \mid X_n = i] = P[X_1 = j \mid X_0 = i] \quad (2)$$

per a qualsevol $n \in \{0, 1, \dots\}$. És a dir, suposem que la cadena de Markov és homogènia en el temps, perquè la probabilitat condicional no depèn del temps n . Les p_{ij} , j , per a $i, j \in S$, s'anomenen probabilitats de transició d'un pas de la cadena de Markov, i la matriu $\mathbf{P} := (p_{ij})_{i,j \in S}$ és la matriu de transició.

Sigui T_i el nombre d'unitats de temps que el procés passa en l'estat i abans de fer un transició a un altre estat. Per la independència, tenim que

$$P[T_i = k] = p_{ii}^{k-1} (1 - p_{ii}), \text{ per a } k = 1, 2, \dots \quad (3)$$

És a dir, T_i té una distribució geomètrica amb paràmetre $p := 1 - p_{ii}$.

Un exemple d'histograma obtingut generant 100 observacions a partir d'una distribució geomètrica amb paràmetre $p = 0,8$ es mostra a la figura 1, en la qual s'ha afegit una funció exponencialment decreixent per mostrar la forma esperada de l'histograma.

Ara, sota certes condicions (que es compliran en les aplicacions considerades en aquest article), podem demostrar que les probabilitats limitants

$$\lim_{n \rightarrow \infty} P[X_n = i] \quad (4)$$

existeixen i es poden obtenir resolent el sistema (vegeu, per exemple, Lefebvre 2007):

$$\tilde{y} = \tilde{y} P, \quad (5)$$

subjecte a

$$\sum_{i=0}^{\infty} \tilde{y} \tilde{y}_i = 1, \quad (6)$$

on

$$\tilde{y} := (\tilde{y}_0, \tilde{y}_1, \dots). \quad (7)$$

Aquestes probabilitats limitants, que també representen la proporció de temps que el procés passa en l'estat i durant un període llarg, quan està en equilibri, ens permetran predir el comportament a llarg termini de la cadena de Markov corresponent. De fet, si el model proposat per a les temperatures mitjanes mensuals i les quantitats mensuals de precipitació a Itàlia es considera realista, aleshores, calculant les probabilitats limitants \tilde{y}_i , podem determinar què hauria de passar després d'un gran nombre de mesos quan el procés s'estabilitzi. Per descomptat, les previsions es basaran en les observacions del conjunt de dades que es van utilitzar per calcular aquestes probabilitats limitants. El model s'hauria d'actualitzar i el \tilde{y}_i s'hauria de recalculer quan hi hagi noves observacions disponibles. A la següent secció, s'implementarà el model anterior en el cas de les temperatures mitjanes mensuals a Itàlia.

3. Temperatura mitjana mensual

Sigui D_n la diferència entre la temperatura mitjana (en graus Celsius) d'un mes determinat i el mes corresponent de l'any anterior. Definim

$$X_n = \begin{cases} 0 & \text{si } D_n \leq \tilde{y}_1, \\ 1 & \text{si } D_n \in (\tilde{y}_1, 0), \\ 2 & \text{si } D_n \in (0, 1), \\ \dots & \text{si } D_n \in [0, 1), \tilde{y}_2 \leq D_n < \tilde{y}_3 \\ & \dots \\ & \text{si } D_n \leq \tilde{y}_1. \end{cases} \quad (8)$$

Primer, utilitzant les dades de <http://sdwebx.worldbank.org/climateportal> Per al període 1991-2015, vam obtenir els histogrames de les variables aleatòries T_i , per a $i = 0, 1, 2, 3$. Es mostren a les figures 2-5. Veiem que els diversos histogrames tenen efectivament la forma esperada per a una distribució geomètrica. Per tant, una cadena de Markov amb els estats anteriors és un model realista per a les variacions de la temperatura mitjana mensual a Itàlia durant el període considerat, cosa que implica que les probabilitats límit \tilde{y}_i es poden utilitzar per predir les variacions de les temperatures mitjanes mensuals després d'un gran nombre de mesos.

Observació. Podríem realitzar la prova estadística de bondat d'ajust de Pearson per comprovar si la distribució geomètrica és un bon model per a les variables aleatòries T_i . Tanmateix, aquí el nombre de graus de llibertat seria molt petit.

A continuació, vam estimar les probabilitats de transició p_{ij} . La matriu de transició estimada és

$$P = \begin{matrix} & \begin{matrix} \tilde{y} & \tilde{y} & \tilde{y} & \tilde{y} \end{matrix} \\ \begin{matrix} \tilde{y} \\ \tilde{y} \\ \tilde{y} \\ \tilde{y} \end{matrix} & \begin{pmatrix} 5/67 & 12/67 & 11/67 & 39/67 \\ 24/69 & 17/69 & 16/64 & 20/64 \\ 37/76 & 22/76 & 11/76 & 6/76 \end{pmatrix} \end{matrix} \quad (9)$$

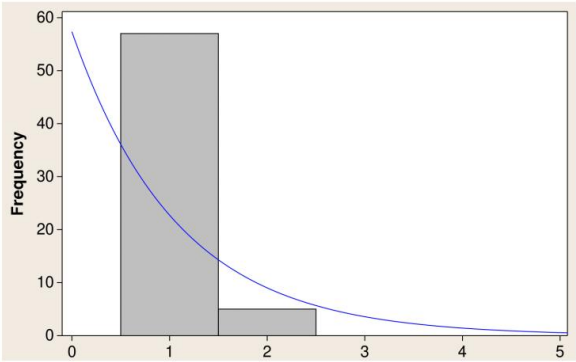


FIGURA 2. Histograma de la variable aleatòria T0 en el cas de la temperatura.

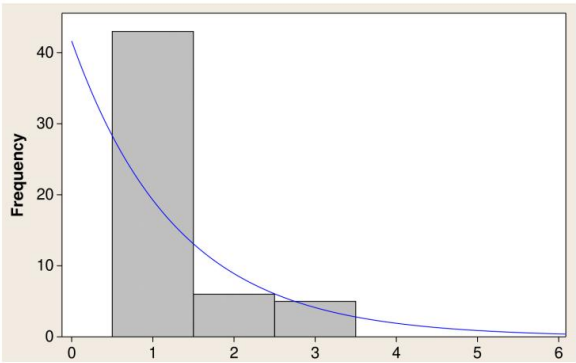


FIGURA 3. Histograma de la variable aleatòria T1 en el cas de la temperatura.

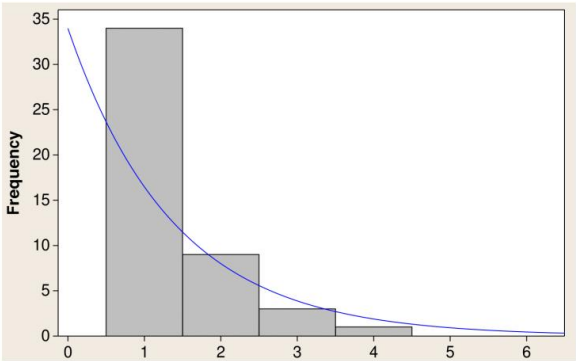


FIGURA 4. Histograma de la variable aleatòria T2 en el cas de la temperatura.

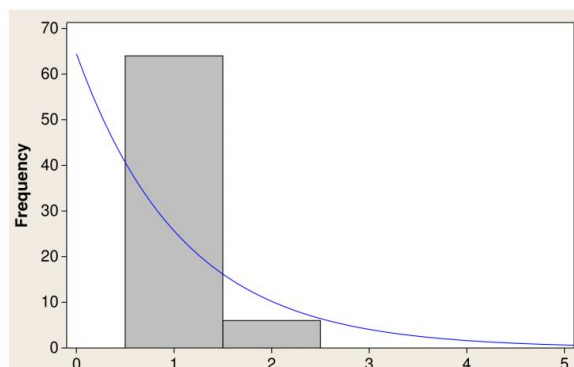


FIGURA 5. Histograma de la variable aleatòria T3 en el cas de la temperatura.

Com que els elements de la matriu són tots diferents de zero, es dedueix immediatament que es compleixen totes les condicions necessàries per a l'existència de les probabilitats límit. Per tant, aquestes probabilitats límit es poden obtenir resolent (5), (6). Trobem que

$$\bar{y}_0 = 0,2512, \bar{y}_1 = 0,2524, \bar{y}_2 = 0,2245, \bar{y}_3 = 0,2719. \quad (10)$$

Cal tenir en compte que les quatre probabilitats límit no són gaire diferents, amb una probabilitat lleugerament més gran de tenir un augment d'almenys 1 grau. La probabilitat de tenir una disminució (0,5036) és gairebé la mateixa que la de tenir un augment (0,4964) de la temperatura mitjana mensual. De fet, la mitjana de les 288 diferències mensuals és aproximadament de 0,065, cosa que és força petita.

Observació. Per definició, les \bar{y}_i són probabilitats limitants. També es poden obtenir calculant **que P** hauria de ser gairebé idèntica. Per a n prou gran, les quatre línies de la matriu **P** s'aproximen a \bar{y} . Aquí, trobem que

$$P_{y16} \begin{pmatrix} \bar{y} \\ \bar{y} \\ \bar{y} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} 0,2512 & 0,2524 & 0,2245 & 0,2719 \\ 0,2245 & 0,2719 & 0,2512 & 0,2524 \\ 0,2512 & 0,2524 & 0,2245 & 0,2719 \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{y} \\ \bar{y} \\ \bar{y} \end{pmatrix}. \quad (11)$$

Per tant, després de 16 mesos les probabilitats dels quatre estats possibles de la cadena de Markov ja són iguals a les probabilitats límit.

Per comprovar si les conclusions són les mateixes si només considerem els anys més recents, vam dividir el conjunt de dades en dues parts iguals (des del principi de 1991 fins al juny de 2003, i des del juliol de 2003 fins al final de 2015) i vam calcular les probabilitats límit per a cada part. Els resultats es mostren a la Taula 1.

Notem que \bar{y}_2 i \bar{y}_3 no van canviar gaire entre el principi i el final del període considerat, però que les probabilitats límit \bar{y}_0 i \bar{y}_1 s'han invertit gairebé, cosa que denota una probabilitat més gran de grans disminucions. La probabilitat de grans canvis positius o negatius ha augmentat d'aproximadament 0,49 a més de 0,55. Per tant, es pot concloure que el canvi climàtic sembla ser observable a Itàlia. A la següent secció, parlarem de les precipitacions mitjanes mensuals a Itàlia.

TAULA 1. Probabilitats límit per als períodes 1991-2015, 1991-2003 i 2003-2015 en el cas de la temperatura.

| Període | p0 | p1 | p2 | p3 |
|-----------|--------|--------|--------|--------|
| 1991-2015 | 0,2512 | 0,2524 | 0,2245 | 0,2719 |
| 1991-2003 | 0,2227 | 0,2884 | 0,2171 | 0,2718 |
| 2003-2015 | 0,2821 | 0,2155 | 0,2292 | 0,2732 |

4. Precipitació mensual mitjana

Sigui ara D_n la diferència entre la pluja mitjana (en mm) per a un mes determinat i el mes corresponent de l'any anterior. Establim

$$X_n = \begin{cases} 0 & \text{si } D_n \in [0, 30], \\ 1 & \text{si } D_n \in (30, 60], \\ 2 & \text{si } D_n \in (60, 90], \\ 3 & \text{si } D_n \in (90, 120]. \end{cases} \tag{12}$$

Pel que fa al cas de la temperatura mitjana mensual, primer observem els histogrames de les variables aleatòries T_i , per a $i = 0, 1, 2, 3$; vegeu les figures 6-9. De nou, podem observar que els diversos histogrames semblen confirmar el fet que les variables aleatòries corresponents tenen aproximadament una distribució geomètrica, tal com es requereix perquè el model sigui realista. Les probabilitats de

$$P = \begin{matrix} & \text{transició} & \text{estimades } p_i, j \text{ són les} \\ \begin{matrix} \vdots \\ \ddots \\ \vdots \end{matrix} & \begin{matrix} \vdots \\ \ddots \\ \vdots \end{matrix} & \begin{matrix} \vdots \\ \ddots \\ \vdots \end{matrix} \\ \begin{matrix} \vdots \\ \ddots \\ \vdots \end{matrix} & \begin{matrix} \vdots \\ \ddots \\ \vdots \end{matrix} & \begin{matrix} \vdots \\ \ddots \\ \vdots \end{matrix} \end{matrix} \tag{13}$$

i les probabilitats límit vénen donades per

$$\ddot{y}_0 = 0,2122, \ddot{y}_1 = 0,2897, \ddot{y}_2 = 0,3097, \ddot{y}_3 = 0,1884. \tag{14}$$

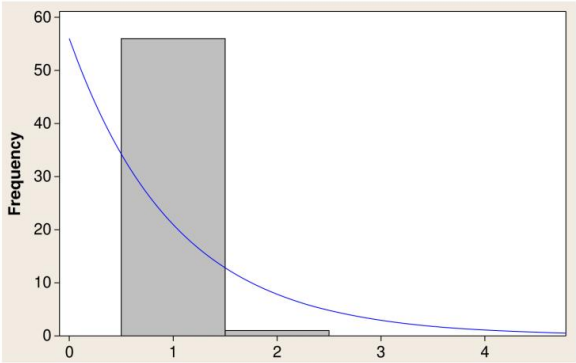


FIGURA 6. Histograma de la variable aleatòria T_0 en el cas de la precipitació.

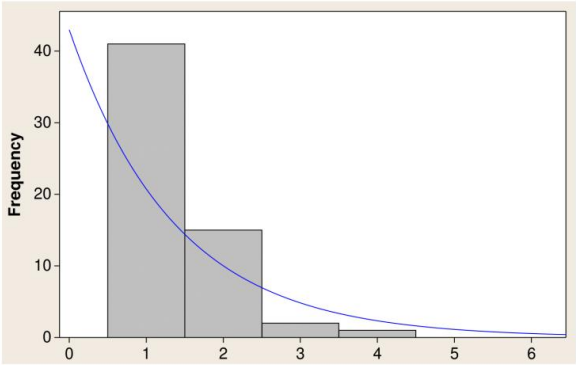


FIGURA 7. Histograma de la variable aleatòria T1 en el cas de la precipitació.

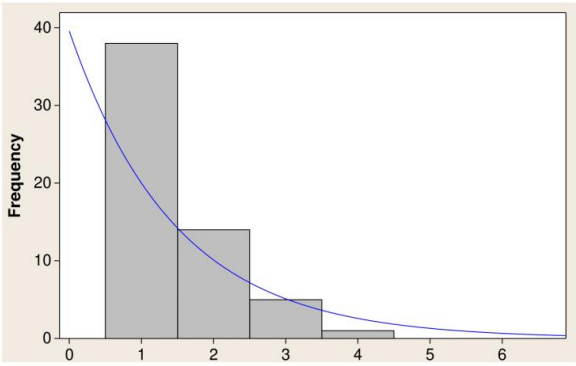


FIGURA 8. Histograma de la variable aleatòria T2 en el cas de la precipitació.

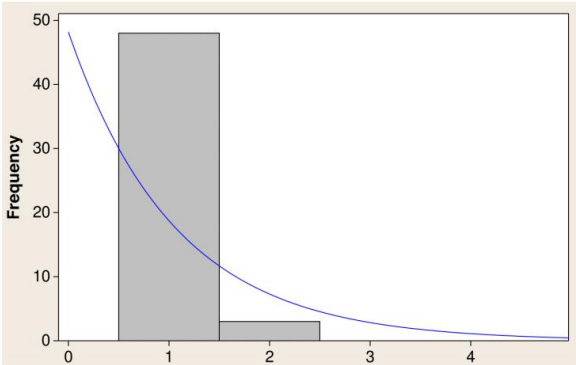


FIGURA 9. Histograma de la variable aleatòria T3 en el cas de la precipitació.

Observació. Com en el cas de la temperatura, trobem que les quatre files de la matriu **P** són iguals a les probabilitats anteriors. D'això es ¹⁶ són dedueix que les previsions que fem són vàlides després d'un màxim de 16 mesos.

Aquesta vegada, tenim molta més variació en les probabilitats límit i una probabilitat menor (0,4006) de grans canvis (amb l'elecció de 30 mm en la definició dels diversos estats). Tanmateix, la probabilitat de tenir una disminució (0,5019) és de nou gairebé la mateixa que la de tenir un augment (0,4981) de la pluja mitjana mensual. La mitjana de les 288 diferències mensuals és d'aproximadament -0,36 mm, cosa que apunta cap a una petita disminució a llarg termini.

Finalment, vam realitzar la mateixa anàlisi amb els dos subconjunts de dades de la mateixa mida. Els resultats, presentats a la Taula 2, són força similars, excepte per un gran augment en el valor de \hat{y}_3 . Els valors de \hat{y}_0 i \hat{y}_1 són força constants, però \hat{y}_2 ha disminuït i \hat{y}_3 ha augmentat.

TAULA 2. Probabilitats límit per als períodes 1991-2015, 1991-2003 i 2003-2015 en el cas de la pluja.

| Període | p0 | <small>página 1</small> | p2 | p3 |
|-----------|--------|-------------------------|--------|--------|
| 1991-2015 | 0,2122 | 0,2897 | 0,3097 | 0,1884 |
| 1991-2003 | 0,2088 | 0,2965 | 0,3278 | 0,1668 |
| 2003-2015 | 0,2155 | 0,2829 | 0,2914 | 0,2102 |

significativament. A més, hi ha hagut un augment d'una probabilitat de 0,38 de grans canvis a 0,43. Sembla que s'ha produït un canvi de règim si observem les mitjanes de les 144 diferències mensuals de cada subconjunt: respectivament +0,11 i $\hat{y}_0,83$. Per tant, el canvi climàtic crea encara més variacions en la pluja que en la temperatura.

5. Conclusió

En aquest article, hem estudiat les temperatures i precipitacions mitjanes mensuals a Itàlia durant el període 1991-2015. Hem vist que, en cada cas, una cadena de Markov de temps discret amb quatre estats possibles és un model adequat per a les dades. Basant-nos en els nostres models, hem pogut estimar el comportament a llarg termini dels canvis de temperatura i quantitat de precipitació.

Les dades semblen confirmar els efectes del canvi climàtic tant en la temperatura com en la pluja a Itàlia, i les precipitacions es veuen especialment afectades.

Agraïments

Aquest treball ha estat finançat pel Consell de Recerca en Ciències Naturals i Enginyeria de Canadà. L'autor també vol agrair al revisor els seus comentaris constructius.

Referències

Avilés, A., Célleri, R., Solera, A., i Paredes, J. (2016). "Predicció probabilística d'esdeveniments de sequera mitjançant models basats en cadenes de Markov i xarxes bayesianes: un estudi de cas d'una conca fluvial regulada als Andes". Water 8(2), 37. DOI: [10.3390/w8020037](https://doi.org/10.3390/w8020037).

- Caskey Jr., JE (1963). «Un model de cadena de Markov per a la probabilitat d'ocurrència de precipitacions en intervals de diverses longituds». *Monthly Weather Review* 91(6), 298–301. DOI: [10.1175/1520-0493\(1963\)091%3c0298:AMCMFT%3e2.3.CO;2](https://doi.org/10.1175/1520-0493(1963)091%3c0298:AMCMFT%3e2.3.CO;2).
- Drton, M., Marzban, C., Guttorp, P., i Schaefer, JT (2003). «Un model de cadena de Markov de l'activitat tornàdica». *Monthly Weather Review* 131(12), 2941–2953. DOI: [10.1175/1520-0493\(2003\)131%3c2941:AMCMOT%3e2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131%3c2941:AMCMOT%3e2.0.CO;2).
- Lefebvre, M. (2007). *Processos estocàstics aplicats*. Nova York, Estats Units: Springer-Verlag. DOI: [10.1007/978-0-387-48976-6](https://doi.org/10.1007/978-0-387-48976-6).
- Lefebvre, M. i Guilbault, J.-L. (2008). «Ús de processos de Poisson filtrats per modelar el flux d'un riu». *Modelització matemàtica aplicada* 32(12), 2792–2805. DOI: [10.1016/j.apm.2007.09.035](https://doi.org/10.1016/j.apm.2007.09.035).
- Matis, JH, Birkett, T., i Boudreaux, D. (1989). «Una aplicació de l'enfocament de la cadena de Markov per a la predicció del rendiment del cotó a partir d'enquestes». *Agricultural Systems* 29(4), 357–370. DOI: [10.1016/0308-521X\(89\)90097-8](https://doi.org/10.1016/0308-521X(89)90097-8).
- Tagliaferri, F., Hayes, BP, Viola, IM, i Djokic, SZ (2016). «Modelització del vent amb cadenes de Markov imbricades». *Journal of Wind Engineering and Industrial Aerodynamics* 157, 118–124. DOI: [10.1016/j.jweia.2016.08.009](https://doi.org/10.1016/j.jweia.2016.08.009).

* Polytechnique Montréal
 Departament de Matemàtiques i Enginyeria Industrial CP
 6079, Succursale Centre-ville, Mont-real, Québec, Canadà H3C 3A7

Correu electrònic: mario.lefebvre@polymtl.ca

