

**POLITECNICO DI MILANO**

Facoltà di Ingegneria Civile, Ambientale e Territoriale

Corso di Laurea Specialistica in Ingegneria per l' Ambiente e il Territorio



**SVILUPPO E ANALISI DI UN MODELLO SPAZIALMENTE  
ESPLICITO DI TRASMISSIONE DEL COLERA INTEGRATO  
CON PREVISIONI DI PRECIPITAZIONE:  
APPLICAZIONE ALL'EPIDEMIA DI HAITI**

Relatore: Prof. Marino GATTO

Correlatore: Ing. Lorenzo RIGHETTO

Tesi di Laurea di:

Barbara CASARTELLI Matr. 751077

Anno accademico 2010-2011

# SOMMARIO<sup>1</sup>

In questo lavoro si propone un modello epidemiologico per descrivere l'epidemia di colera diffusasi ad Haiti da ottobre 2010 e tuttora in corso. In particolare, si utilizza un modello Suscettibili-Infetti-Ristabiliti-Batteri spazialmente esplicito, che descrive la diffusione di *Vibrio cholerae* attraverso il trasporto idrologico e gli spostamenti sia degli individui suscettibili che di quelli infetti tra le diverse comunità. Dopo aver osservato una relazione tra la risorgenza dell'epidemia e l'inizio della stagione delle piogge, si propone una variante al modello che considera esplicitamente la precipitazione come una forzante esterna, che modifica il tasso di contaminazione da *Vibrio cholerae*. L'obiettivo di questo lavoro è quello di effettuare delle previsioni di precipitazione da inserire come ingresso nel modello epidemiologico suddetto, per poter generare degli scenari di diffusione dell'epidemia. Si propongono diverse alternative per prevedere le precipitazioni, in particolare si analizzano metodi che si basano sulle statistiche della serie storica (media, 25-esimo percentile e 75-esimo percentile), sulle precipitazioni dell'anno tipico, su modelli ARX e su generatori puntuali di precipitazione poissoniani. Il modello, che è stato calibrato fino al 28 maggio 2011, viene utilizzato per simulare l'evoluzione dell'epidemia di colera fino a dicembre 2013. Si mostra che, con l'introduzione di previsioni di precipitazione, è possibile predire anche l'andamento dei casi di colera che, nella fattispecie, aumentano in corrispondenza della stagione piovosa verificatasi tra giugno e settembre 2011. Il modello prevede inoltre delle risorgenze della malattia in corrispondenza di settembre 2012 e di settembre 2013. Il valore intrinseco di questa analisi risiede principalmente nel gran numero di possibili scenari di precipitazione generati e, di conseguenza, anche epidemiologici, che permettono di implementare delle politiche di intervento adeguate allo sviluppo previsto della malattia, sia nel breve che nel lungo periodo.

---

<sup>1</sup> In fondo al testo è inserito un riassunto più esteso che ripercorre i passaggi fondamentali e i risultati ottenuti dalla presente tesi.

# ABSTRACT

In this work we propose an epidemiological model to study the current Haiti cholera outbreak, began in October 2010. In particular, the spread of *Vibrio cholerae* is described via a spatially explicit Susceptible-Infected-Recovered-Bacteria model, accounting for hydrological transport and for the mobility of susceptible and infected individuals among different communities. Given the correspondence between the epidemic resurgence and the beginning of rainy season, we propose a variant of the model to explicitly account for rainfall as a driver increasing the baseline contamination rate. The aim of this work is to generate rainfall forecasts to utilize them as an input of the epidemiological model. In this way it is possible to study different scenarios for the evolution of the disease. We propose different options to perform rainfall forecasts, specifically we analyze methods based on statistical properties of the time series (mean, 25<sup>th</sup> percentile and 75<sup>th</sup> percentile), on typical year's rainfall, on ARX models and on point rainfall Poisson's generators. The epidemiological model, calibrated until 28<sup>th</sup> May 2011, is used to predict the evolution of the cholera outbreak up to December 2013. We show that, by introducing precipitation as an epidemiological driver, it is possible to describe also the observed pattern of cholera cases until September 2011, which in particular increase in correspondence of the rainy season, starting in June. Moreover, when used as a predictor the model shows a resurgence of cholera in September 2012 and in September 2013. The intrinsic value of this analysis consists in the generation of a large number of possible rainfall scenarios and therefore also epidemiological ones, which permit to implement intervention policies and strategies appropriated to the expected development of the disease, both in the short and in the long term.

# INDICE

<b>SOMMARIO</b> .....	<b>II</b>
<b>ABSTRACT</b> .....	<b>III</b>
<b>INDICE</b> .....	<b>IV</b>
<b>INDICE DELLE FIGURE</b> .....	<b>VI</b>
<b>INTRODUZIONE</b> .....	<b>1</b>
<b>1 INQUADRAMENTO DEL CASO DI STUDIO</b> .....	<b>4</b>
1.1    IL COLERA.....	5
1.2    HAITI .....	8
1.2.1    Economia .....	8
1.2.2    Demografia .....	9
1.3    TERREMOTO AD HAITI .....	9
1.4    L'EPIDEMIA DI COLERA .....	12
1.5    MODELLI EPIDEMIOLOGICI ESISTENTI .....	15
<b>2 MODELLO EPIDEMIOLOGICO</b> .....	<b>18</b>
2.1    IL MODELLO SIRB (SUSCETTIBILI-INFETTI-RISTABILITI-BATTERI) .....	19
2.1.1    Il modello epidemiologico locale .....	19
2.1.2    Il modello spazialmente esplicito .....	23
2.2    IL MODELLO SIRB CON LE PRECIPITAZIONI.....	28
2.3    STIMA E CALIBRAZIONE DEI PARAMETRI.....	30
<b>3 DATI</b> .....	<b>34</b>
3.1    DATI EPIDEMIOLOGICI .....	35
3.2    DATI RELATIVI ALLA POPOLAZIONE .....	35
3.3    DATI UTILIZZATI PER L'ESTRAZIONE DELLA RETE IDROLOGICA .....	36
3.4    PRECIPITAZIONI .....	38
3.4.1    Statistiche di precipitazione .....	39
3.5    CORRELAZIONE SPAZIALE DELLA PRECIPITAZIONE .....	43

<b>4 VALIDAZIONE E PREVISIONE.....</b>	<b>46</b>
4.1 VALIDAZIONE CON PRECIPITAZIONI REALI .....	47
4.2 SIMULAZIONE CON L'ANNO DI PRECIPITAZIONE MEDIO .....	49
4.3 SIMULAZIONE CON LA PRECIPITAZIONE DELL'ANNO TIPICO .....	54
4.4 SIMULAZIONE CON LA PRECIPITAZIONE GENERATA CON UN GENERATORE PUNTUALE DI PRECIPITAZIONE POISSONIANO .....	56
4.4.1 <i>Generazione per dipartimenti disaggregati</i> .....	60
4.4.2 <i>Generazione per dipartimenti aggregati</i> .....	63
4.5 SIMULAZIONE CON LA PRECIPITAZIONE GENERATA CON UN GENERATORE PUNTUALE DI PRECIPITAZIONE POISSONIANO A PARTIRE DAI BACINI .....	68
4.5.1 <i>Generazione di precipitazione partendo da dati giornalieri</i> .....	70
4.5.2 <i>Generazione di precipitazione partendo da dati decadali</i> .....	73
4.6 SIMULAZIONE CON LA PRECIPITAZIONE CALCOLATA CON MODELLI ARX.....	75
4.7 CONFRONTO TRA GLI <i>OUTPUT</i> EPIDEMIOLOGICI OTTENUTI CON I DIVERSI METODI DI PREVISIONE DELLA PRECIPITAZIONE.....	83
<b>5 CONCLUSIONI .....</b>	<b>85</b>
5.1 SVILUPPI FUTURI.....	89
<b>6 BIBLIOGRAFIA.....</b>	<b>90</b>
<b>RIASSUNTO .....</b>	<b>97</b>

# INDICE DELLE FIGURE

<b>Figura 1.1:</b> Foto al microscopio elettronico del batterio <i>Vibrio cholerae</i> .....	6
<b>Figura 1.2:</b> Mappa globale dei casi di colera registrati tra il 2010 e il 2011 .....	7
<b>Figura 1.3:</b> Demografia divisa per distretti .....	10
<b>Figura 1.4:</b> Localizzazione dell'epicentro del terremoto del 12 gennaio 2010.....	11
<b>Figura 1.5:</b> Grafico dei casi di colera per ogni dipartimento.....	14
<b>Figura 2.1:</b> Schema del modello epidemiologico locale. ....	20
<b>Figura 2.2:</b> Rappresentazione grafica del modello spazialmente esplicito.....	23
<b>Figura 2.3:</b> Modellizzazione del trasporto idrologico .....	24
<b>Figura 2.4:</b> Modellizzazione del trasporto dovuto alla mobilità umana .....	27
<b>Figura 2.5:</b> Relazione tra i casi di colera e la precipitazione.....	29
<b>Figura 3.1:</b> Mappa dei bacini idrologici iniziali.....	37
<b>Figura 3.2:</b> Mappa relativa ai 366 bacini finali .....	37
<b>Figura 3.3:</b> Dati di precipitazione .....	39
<b>Figura 3.4:</b> Andamento delle precipitazioni giornaliere e di quelle decadali.....	40
<b>Figura 3.5:</b> Media annuale della precipitazione di Haiti per gli anni dal 2002 al 2010.....	40
<b>Figura 3.6:</b> Andamento medio annuale della precipitazione di Haiti.....	41
<b>Figura 3.7:</b> Andamento della precipitazione media dei dieci distretti di Haiti.....	42
<b>Figura 3.8:</b> Mappe di correlazione della precipitazione .....	44
<b>Figura 4.1:</b> Calibrazione e validazione del modello epidemiologico. ....	48
<b>Figura 4.2:</b> Simulazione con l'anno di precipitazione medio.....	50
<b>Figura 4.3:</b> Casi settimanali di colera per ogni distretto simulati con l'anno di precipitazione medio.....	52
<b>Figura 4.4:</b> Casi cumulati di colera per ogni distretto simulati con l'anno di precipitazione medio .....	53
<b>Figura 4.5:</b> Simulazione con la precipitazione dell'anno tipico. ....	55
<b>Figura 4.6:</b> Schema delle grandezze caratteristiche in un processo di tipo poissoniano. ....	57
<b>Figura 4.7:</b> Statistiche calcolate sui dati generati con il generatore puntuale di precipitazione di Poisson per il distretto di Nord-Ouest .....	59

<b>Figura 4.8:</b> Simulazione effettuata con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione dei dipartimenti disaggregati. ....	61
<b>Figura 4.9:</b> Casi di colera settimanali per ogni distretto simulati con la precipitazione generata con il generatore puntuale di precipitazione di Poisson partendo dalla precipitazione giornaliera dei distretti disaggregati .....	62
<b>Figura 4.10:</b> Suddivisione di Haiti in zone con precipitazioni comparabili. ....	64
<b>Figura 4.11:</b> Andamento delle precipitazioni medie mensili dei distretti di Nord e Nord-Est e di Artibonite e Ouest .....	65
<b>Figura 4.12:</b> Andamento delle precipitazioni medie mensili dei distretti di Grande Anse, Nippes, Sud e Sud-Est .....	65
<b>Figura 4.13:</b> Simulazione con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei dipartimenti aggregati .....	67
<b>Figura 4.14:</b> Casi di colera settimanali per ogni distretto simulati con la precipitazione generata con il generatore puntuale di precipitazione di Poisson partendo dalla precipitazione giornaliera dei distretti aggregati.....	68
<b>Figura 4.15:</b> <i>Downscaling</i> della precipitazione dai dipartimenti ai bacini .....	69
<b>Figura 4.16:</b> Simulazione con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei nodi. ....	72
<b>Figura 4.17:</b> Simulazione con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione decadale dei nodi. ....	74
<b>Figura 4.18:</b> Precipitazione simulata con i tre migliori modelli ARX per il distretto Nord-Ovest 80	
<b>Figura 4.19:</b> Simulazione dei casi di colera per tutta Haiti utilizzando i modelli ARX.....	81
<b>Figura 4.20:</b> Zoom temporale sulla simulazione dei casi di colera simulati basandosi sulla precipitazione generata utilizzando i modelli ARX.....	82
<b>Figura 5.1:</b> Simulazione con l'anno di precipitazione medio confrontato con i dati reali aggiornati al 20 gennaio 2012. ....	88
<b>Figura 1:</b> Relazione tra i casi di colera e la precipitazione.....	120
<b>Figura 2:</b> Rappresentazione dei dati di precipitazione .....	120
<b>Figura 3:</b> Media dell'andamento annuale della precipitazione di Haiti .....	121
<b>Figura 4:</b> Calibrazione e validazione del modello epidemiologico .....	122
<b>Figura 5:</b> Simulazione con l'anno di precipitazione medio.....	123
<b>Figura 6:</b> Simulazione con la precipitazione dell'anno tipico .....	124
<b>Figura 7:</b> Statistiche calcolate sui dati generati con il generatore puntuale di precipitazione di Poisson per il distretto di Nord-Ouest .....	125
<b>Figura 8:</b> Simulazione con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei dipartimenti disaggregati.....	126
<b>Figura 9:</b> Mappe di correlazione della precipitazione .....	127

<b>Figura 10:</b> Suddivisione di Haiti in zone con precipitazioni comparabili .....	128
<b>Figura 11:</b> Andamento delle precipitazioni medie mensili dei distretti di Nord e Nord-Est e di Artibonite e Ouest .....	128
<b>Figura 12:</b> Andamento delle precipitazioni medie mensili dei distretti di Grande Anse, Nippes, Sud e Sud-Est .....	129
<b>Figura 13:</b> Simulazione con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei dipartimenti aggregati .....	130
<b>Figura 14:</b> Casi di colera settimanali per ogni distretto simulati con la precipitazione generata con il generatore di precipitazione puntuale di Poisson partendo dalla precipitazione giornaliera dei distretti disaggregati .....	131
<b>Figura 15:</b> Casi di colera settimanali per ogni distretto simulati con la precipitazione generata con il generatore puntuale di precipitazione di Poisson partendo dalla precipitazione giornaliera dei distretti aggregati.....	132
<b>Figura 16:</b> Simulazione con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei nodi. ....	133
<b>Figura 17:</b> Simulazione con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione decadale dei nodi.....	134
<b>Figura 18:</b> Precipitazione simulata con i modelli ARX per il distretto Nord-Ovest .....	135
<b>Figura 19:</b> Zoom temporale sulla simulazione dei casi di colera simulati basandosi sulla precipitazione generata utilizzando i modelli ARX.....	135
<b>Figura 20:</b> Simulazione con l'anno di precipitazione medio confrontato con i dati reali aggiornati al 20 gennaio 2012. ....	136



## INTRODUZIONE

Secondo la *World Health Organization* (WHO), le malattie diarroiche sono la quinta causa di morte al mondo e causano più di 2 milioni di decessi l'anno (3.7% dei decessi totali) [WHO, 2008]. Il colera, probabilmente la più conosciuta tra le malattie di questo tipo, è responsabile ogni anno in media di 4 000-6 300 decessi e colpisce 178 000-327 000 persone. Tuttavia, spesso i casi imputabili al colera non sono identificati e vengono classificati come diarrea acuta. Infatti la WHO stima che il numero reale dei casi possa aggirarsi intorno ai 3-5 milioni e quello dei decessi intorno a 100 000-130 000 all'anno [WHO, 2010a]. Dal 2009 al 2010 i casi registrati sono aumentati del 43% a livello globale e del 130% rispetto al 2000 [WHO, 2011].

La più recente epidemia è stata registrata ad Haiti nell'ottobre 2010 e ha avuto il suo massimo a gennaio 2011. A causa del terremoto di magnitudo 7.0  $M_w$  che ha colpito Haiti il 12 gennaio 2010 (il cui epicentro è stato localizzato a circa 25 km dalla capitale Port-au-Prince), la maggior parte delle infrastrutture, incluse quelle idriche, è andata distrutta. Questa situazione ha favorito la diffusione del colera, malattia principalmente collegata a un difficile accesso all'acqua potabile e a una carenza delle condizioni igienico-sanitarie. Il primo caso di colera registrato ad Haiti risale al 21 ottobre 2010, vicino a Mirebalais, nel distretto di Centre [Piarroux *et al.*, 2011] e in prossimità del fiume Artibonite, il più grande di Haiti, che percorre il territorio haitiano attraversando non solo il distretto di Centre, ma anche quello di Artibonite. A metà gennaio 2012, sono stati osservati 520 000 casi di colera, dei quali 7 005 hanno portato al decesso [PAHO, 2012].

Il colera, che è causato dal batterio *Vibrio cholerae*, è una malattia che si trasmette tramite l'acqua. Una volta nell'intestino il batterio trova le condizioni ideali per potersi riprodurre, inducendo diarrea e vomito nell'individuo ospite. Questi sintomi possono causare una severa disidratazione e, nei casi peggiori, la morte. Generalmente il colera può essere trattato con successo attraverso la reidratazione orale o, nei casi più gravi, per via endovenosa.

Il tempo di incubazione del colera varia dalle 2 ore ai 5 giorni, il che in parte spiega l'andamento esplosivo delle epidemie [WHO, 2010b]. In molti casi gli infetti non mostrano alcun sintomo (circa l'80%). I batteri sono presenti nelle feci dei malati (sintomatici e asintomatici) e possono quindi essere espulsi nell'ambiente esterno, che viene così contaminato.

La veloce diffusione della malattia e il numero di persone contagiate mostrano l'importanza di sviluppare un metodo adeguato in grado di prevedere e ridurre l'impatto dell'epidemia. A questo scopo è necessario implementare un opportuno modello quantitativo che permetta l'analisi preventiva dell'impatto di politiche di pubblica sanità.

Poiché è stata osservata una relazione tra l'aumento dei casi di infetti e le precipitazioni (infatti la risorgenza del colera è coincisa proprio con la stagione delle piogge), nella presente tesi viene validato un modello epidemiologico che considera le precipitazioni come una forzante esterna che va ad incidere sulla severità del colera a livello locale. Di conseguenza, per prevedere l'evoluzione dell'epidemia di colera che sta colpendo Haiti dal 2010, è necessario eseguire delle previsioni di precipitazione, da inserire come ingresso nel modello usato per le simulazioni su breve e lungo periodo. Vengono proposti diversi metodi per generare la precipitazione e si confrontano i diversi scenari ottenuti.

Dopo aver analizzato la dinamica dell'epidemia (**Capitolo 1**), si descrive il modello scelto per le simulazioni. Per la precisione, si utilizza un modello epidemiologico SIRB (Suscellibili-Infetti-Ristabiliti-Batteri) che introduce l'effetto delle precipitazioni sulla diffusione della malattia (**Capitolo 2**).

I dati per tarare e validare il modello sono reperiti dalla *Pan American Health Organization* (PAHO<sup>2</sup>), mentre, per quanto riguarda la precipitazione, dalla *NASA-JAXA's Tropical Rainfall Measuring Mission* e dall'*International Research Institute for Climate and Society* (**Capitolo 3**).

Poiché si vogliono effettuare delle previsioni sulla propagazione dell'epidemia, è necessario avere a disposizione delle previsioni di precipitazione. Sarà questo un argomento particolarmente importante ai fini di questa tesi, in quanto verranno discussi i risultati ottenuti utilizzando diversi metodi di previsione delle precipitazioni e i conseguenti scenari epidemiologici ottenuti (**Capitolo 4**).

Nel **Capitolo 5** si propongono delle possibili varianti al modello di colera e ai metodi di previsione delle precipitazioni, che potranno essere sviluppati in futuro.

Infine si allega un riassunto della presente tesi, che ripercorre tutti i passaggi fondamentali e i risultati ottenuti (**Riassunto**).

---

<sup>2</sup> La PAHO è un ente sanitario internazionale che ha lo scopo di migliorare gli standard sanitari delle nazioni appartenenti sia all'America settentrionale che a quella meridionale. Collabora regolarmente con la *World Health Organization* (WHO). Il sito di riferimento è <http://new.paho.org/>.

## INQUADRAMENTO DEL CASO DI STUDIO

In questo capitolo si descrivono le caratteristiche del batterio *Vibrio cholerae* e le condizioni favorevoli alla sua diffusione. Successivamente si delinea un breve profilo di Haiti, focalizzandosi sulla situazione dopo il terremoto. Infatti, in seguito al sisma di 7.0  $M_w$ , il cui epicentro è stato localizzato a circa 25 km dalla capitale Port-au-Prince, si sono create le condizioni adatte alla diffusione del colera, quali un difficile accesso all'acqua potabile, a causa del severo danneggiamento delle infrastrutture idriche, e una carenza dei servizi igienico-sanitari.

Secondo una stima del *Centers for Disease Control and Prevention*, dal 21 ottobre 2010 (data in cui è stato osservato il primo caso) al 2 settembre 2011 l'epidemia ha coinvolto più del 5% della popolazione, con un tasso di decesso, tra coloro che hanno contratto il colera, dell'1%.

Infine, viene illustrato un quadro generale dei modelli epidemiologici per trattare la diffusione del colera, tra i quali verrà applicato quello di *Rinaldo et al.* (2012) per eseguire le simulazioni sull'andamento dell'epidemia di colera ad Haiti.

## 1.1 Il colera

Il colera è una malattia infettiva del tratto intestinale, che provoca diarrea profusa, spesso complicata con acidosi, ipopotassiemia<sup>3</sup> e vomito. È causato da un batterio Gram-negativo a forma di virgola: il *Vibrio cholerae* (**Figura 1.1**), identificato per la prima volta nel 1859 dall'anatomista italiano Filippo Pacini e studiato dettagliatamente nel 1886 dal medico tedesco Robert Koch. Il nome deriva dal greco *choléra*, da *cholé* che significa "bile".

Le aree dove si manifestano casi clinici di colera a ricorrenza stagionale sono definite aree endemiche (India, Bangladesh e alcune zone dell'Africa). Si conoscono circa 200 sierotipi di vibrione del colera [WHO, 2010b] ma di essi solamente il sierotipo O1 e O139 (quest'ultimo scoperto nel 1992 per un'epidemia avvenuta nel Bangladesh [Comstock et al., 1995]) sono responsabili delle epidemie. L'intestino umano è un ambiente favorevole per la riproduzione di *Vibrio cholerae*, la cui principale fonte di propagazione nell'ambiente sono le feci umane [Sack et al., 2004]. Se in simbiosi con altri microrganismi più complessi, come il fitoplancton e lo zooplancton, il batterio può sopravvivere anche nell'ecosistema acquatico, e quindi all'esterno dell'organismo umano [Colwell, 1996; Islam et al., 1990].

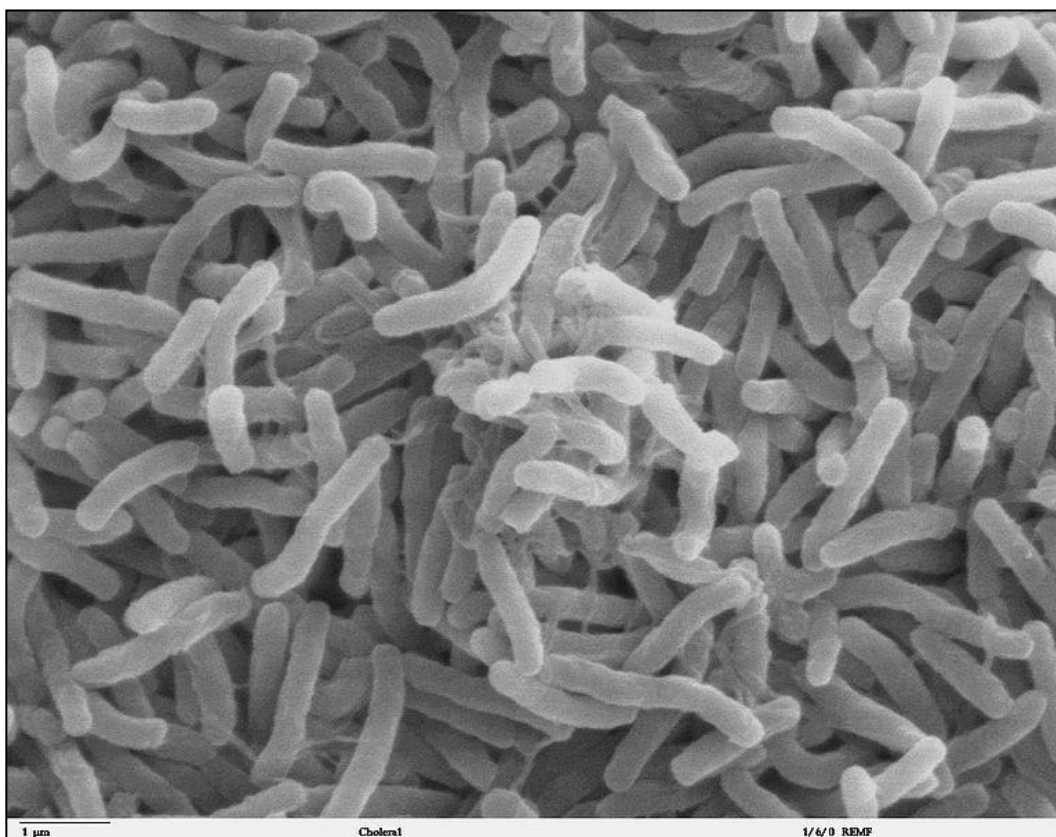
Il batterio si trasmette per via oro-fecale, ovvero quando l'agente patogeno passa, in qualche modo, dall'apparato digerente di un individuo infetto (che eventualmente può essere anche asintomatico) a quello di un altro individuo sano. Tipicamente, questo avviene tramite ingestione di cibo o acqua contaminati.

Esistono due principali vie di contaminazione dell'acqua:

- 1) Presenza del batterio *Vibrio cholerae* nell'ambiente naturale
- 2) Escrezioni di feci contaminate che raggiungono l'ambiente acquatico a causa di un insufficiente trattamento dei reflui.

---

<sup>3</sup> L'ipopotassiemia (o ipokaliemia) è la carenza di potassio nel sangue.



**Figura 1.1:** Foto al microscopio elettronico del batterio *Vibrio cholerae*, che infetta il sistema digerente [Kirn *et al.*, 2000].

Le manifestazioni del colera sono variabili da uno stato asintomatico (circa l'80% dei casi) a uno di diarrea profusa. In quest'ultimo caso si può arrivare fino a una perdita di un litro di feci in un'ora, con conseguente stato di disidratazione che può culminare in uno stato di shock ipovolemico<sup>4</sup>. Deplezioni di liquidi superiori al 10% del peso dell'individuo possono causare il coma e la morte. Il grave stato di disidratazione può determinare la morte nel 30-50% delle persone che non vengono adeguatamente curate.

Lo scopo fondamentale del trattamento del colera è la reidratazione e la correzione degli squilibri elettrolitici. Il metodo più comune è la reidratazione per via orale ma, se i sintomi sono molto severi, potrebbe essere necessaria la reidratazione per via endovenosa. Il trattamento serve per integrare l'acqua e i sali minerali persi

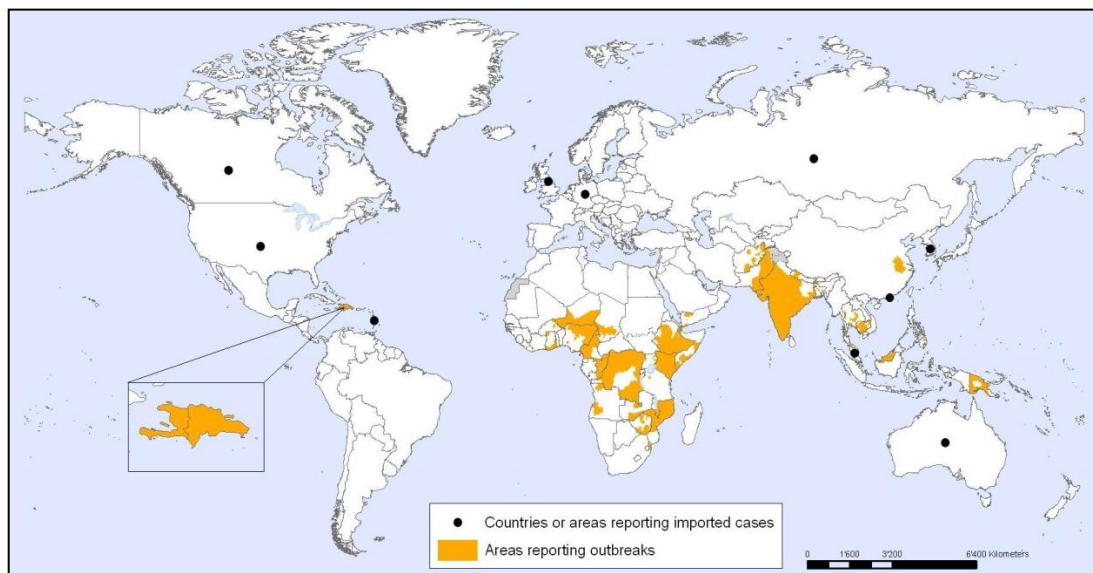
---

<sup>4</sup> È uno stato di shock causato da una diminuzione della massa sanguigna circolante.

(come il magnesio e il sodio). Gli antibiotici possono aiutare la guarigione, diminuendo la durata della malattia e il bisogno di reidratazione. Una volta contratto, il colera lascia gli individui che guariscono in uno stato di immunità, la cui durata è stimata dalle 12 settimane ai 10 anni [King *et al.*, 2008; Koelle *et al.*, 2005].

Per prevenire il contagio è indispensabile migliorare i servizi igienico-sanitari e garantire una buona qualità dell'acqua, attraverso azioni di trattamento e purificazione, ai fini di migliorarne la potabilizzazione. È possibile somministrare il vaccino alla popolazione suscettibile, ma sono necessarie almeno tre settimane perché diventi efficace e la protezione dura per un tempo limitato (in genere qualche anno) [WHO, 2010b].

Recentemente, come mostrato in **Figura 1.2**, si sono registrate epidemie di colera in Bangladesh, Angola, Etiopia, Somalia, Nigeria, Repubblica Democratica del Congo e Zambia [WHO, 2012]. A fine ottobre 2010 il *Centers for Disease Control and Prevention* (CDC) ha confermato che la malattia che colpiva la popolazione haitiana era colera.



**Figura 1.2:** Mappa globale dei casi di colera registrati tra il 2010 e il 2011. In arancione sono indicate le aree in cui sono state registrate delle epidemie, mentre con i pallini neri sono riportati i Paesi in cui sono stati importati dei casi [WHO, 2012].

## 1.2 Haiti

Haiti (ufficialmente Repubblica di Haiti) è una nazione situata nella parte occidentale dell'isola di Hispaniola, nelle Grandi Antille, e confina a est con la Repubblica Dominicana. La superficie di Haiti è di circa 27 750 km<sup>2</sup>. È bagnata a nord dall'Oceano Atlantico, a sud dal Mar dei Caraibi e a ovest dal Canale Sopra Vento, che la separa da Cuba. Haiti è divisa in dieci dipartimenti. La capitale è Port-au-Prince, situata nel distretto di Ouest. Un tempo colonia francese, è stata una delle prime nazioni del continente americano a dichiarare la propria indipendenza (nel 1804).

L'isola ha un territorio prevalentemente montuoso, di origine vulcanica. Le principali catene montuose comprendono il Massif de la Hotte sul promontorio meridionale, con la vetta alta 2347 m del Pic Macaya e il Massif de la Selle, che si estende da ovest a est, immediatamente a sud-est di Port-au-Prince, culminando nel Pic La Selle (2674 m). Al centro sono presenti le Montagnes Noires, la Chaîne des Matheaux e la Chaîne du Trau d'Eau, mentre, a nord, le Montagnes du Nord. Le catene montuose sono separate da profonde vallate e pianure. Lungo i versanti delle montagne scorrono numerosi torrenti, dei quali l'unico navigabile è l'Artibonite, che inizia dal confine dominicano e sfocia immediatamente a nord della città di Saint-Marc.

Il dipartimento di Centre, dove sono stati registrati i primi casi di colera, oltre ad essere attraversato da questo fiume, è anche il distretto in cui sono presenti le catene montuose da cui nascono molti torrenti che si snodano in tutto il territorio haitiano. Pertanto, poiché una delle vie di diffusione del colera è attraverso la rete idrologica, la diffusione della malattia può essere stata favorita dalla posizione geografica del primo focolaio dell'epidemia.

### 1.2.1 Economia

Haiti è il Paese meno sviluppato e più povero del continente americano e dell'intero emisfero occidentale. Secondo il rapporto pubblicato nell'aprile 2011 dall'*International Monetary Fund*, Haiti occupa il 164° posto su 180 classificati in



base al PIL procapite (PPA) [*International Monetary Fund World Economic and Financial Surveys*, 2011].

Circa l'80% della popolazione si trova in una condizione di povertà degradante e il 54% vive con meno di un dollaro al giorno. I disoccupati di Haiti rappresentano oltre il 60% della popolazione e sul paese grava un pesante debito. Quasi il 70% degli haitiani è impiegato nel settore agricolo, che rappresenta quasi un terzo del PIL nonostante sia per lo più una forma di agricoltura di sussistenza praticata su piccola scala. L'industria riveste un ruolo assolutamente marginale.

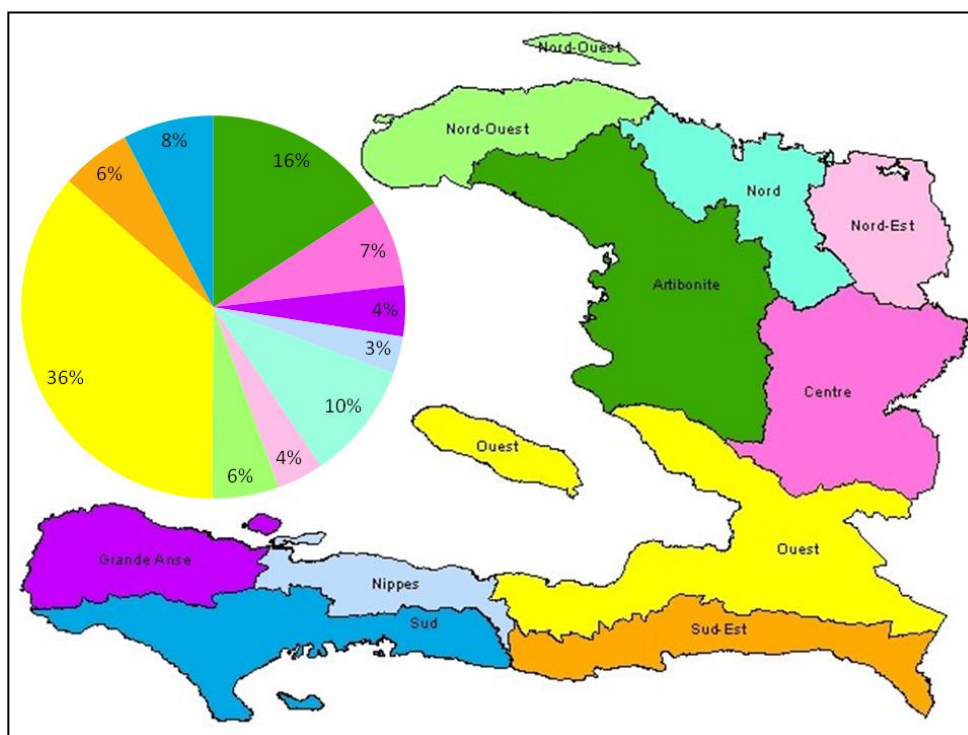
## 1.2.2 Demografia

Haiti ha circa otto milioni di abitanti. Malgrado una densità molto elevata (312 ab/km<sup>2</sup>), la distribuzione della popolazione è fortemente disomogenea: gran parte degli haitiani, infatti, vive nelle città, nelle pianure costiere e nelle valli.

Secondo le stime del *Centers for Disease Control and Prevention* [CDC, 2011], il distretto più popolato è Ouest, che comprende il 36% della popolazione (**Figura 1.3**), la maggior parte della quale risiede nella capitale. Il distretto di Centre, in cui è iniziata l'epidemia, è il dipartimento con la minore densità abitativa (**Tabella 1.1**), ma Artibonite, il distretto a valle di Centre, è invece molto popolato.

## 1.3 Terremoto ad Haiti

Martedì 12 gennaio 2010 un terremoto di magnitudo 7.0  $M_w$  ha colpito Haiti, causando numerose vittime e distruggendo molti edifici. L'epicentro è stato localizzato a circa 25 km in direzione ovest-sud-ovest della città di Port-au-Prince (**Figura 1.4**), che è stata parzialmente distrutta. L'ipocentro è stato individuato a circa 13 km di profondità. La scossa principale si è verificata alle ore 16:53:09 locali (21:53:09 UTC), seguita da un allarme tsunami diramato dal *Pacific Tsunami Warning Center*, che poi è rientrato. Lo *United States Geological Survey* (USGS) ha registrato una lunga serie di repliche nelle prime ore successive al sisma, quattordici delle quali di magnitudo compresa tra 5.0 e 5.9  $M_w$ .



**Figura 1.3:** Nella mappa sono rappresentati i dipartimenti in cui è divisa Haiti. Per ogni dipartimento nel grafico a torta è riportata la percentuale di popolazione corrispondente.

Dipartimento	Capoluogo	Superficie (km <sup>2</sup> )	Popolazione (ab)	Densità abitativa (ab/km <sup>2</sup> )
<b>Artibonite</b>	Gonaïves	4984	1 361 173	273
<b>Centre</b>	Hinche	3675	625 235	170
<b>Grande Anse</b>	Jérémie	1871	368 854	197
<b>Nippes</b>	Miragoâne	1219	270 560	222
<b>Nord</b>	Cap-Haïtien	2106	851 224	404
<b>Nord-Est</b>	Fort-Liberté	1805	337 101	187
<b>Nord-Ouest</b>	Port-de-Paix	2176	482 550	222
<b>Ouest</b>	Port-au-Prince	4827	3 117 344	646
<b>Sud-Est</b>	Jacmel	2023	508 207	251
<b>Sud</b>	Les Cayes	2794	656 513	235
<b>Totale</b>	Port-au-Prince	27480	8 578 761	312

**Tabella 1.1:** Per ogni distretto sono elencati il capoluogo, la superficie, la popolazione e la densità abitativa [CDC, 2010]



**Figura 1.4:** Localizzazione dell'epicentro del terremoto del 12 gennaio 2010 [CIA, 2010].

A causa della povertà e dell'isolamento del Paese e in seguito a gravi danni alle infrastrutture di comunicazione, non è possibile definire con certezza il numero delle vittime del sisma. Secondo una stima dell'ONU al 24 febbraio 2010, il sisma avrebbe provocato 222 517 vittime, 310 900 feriti e oltre 1.5 milioni di sinistrati [ONU, 2010]. La città più colpita è stata la capitale Port-au-Prince (MM<sup>5</sup> IX, secondo la stima USGS), dove molti edifici sono stati gravemente distrutti o danneggiati. Tra questi rientrano il Palazzo Presidenziale, la sede dell'Assemblea Nazionale di Haiti, la cattedrale e la prigione principale.

Tutti gli ospedali della città sono stati distrutti o sono risultati talmente danneggiati da dover essere evacuati. La maggior parte delle infrastrutture è andata

---

<sup>5</sup> Scala Mercalli Modificata che associa l'intensità locale del terremoto a grandezze fisiche misurabili localmente.

distrutta e questo ha inevitabilmente causato ritardi e problemi nella distribuzione degli aiuti umanitari.

Avvertendo immediatamente la gravità della situazione ad Haiti, sono subito arrivati aiuti da parte di numerose nazioni, seguiti da quelli di organizzazioni umanitarie italiane e internazionali. Obiettivi principali in questa prima fase della crisi sono stati garantire l'accesso all'acqua potabile, ripristinare servizi igienici di emergenza, per prevenire le epidemie, identificare e proteggere da abusi e traffico i tanti bambini rimasti orfani o separati dai familiari nel caos della crisi. Ciononostante, dieci mesi dopo la situazione si è aggravata ulteriormente a causa dell'epidemia di colera che si è diffusa in tutta Haiti. Il terremoto ha causato sia una carenza delle condizioni igienico-sanitarie, sia la distruzione o il danneggiamento delle infrastrutture idriche, comportando così un difficile accesso all'acqua potabile e di conseguenza rendendo la popolazione più vulnerabile a questa malattia.

## **1.4 L'epidemia di colera**

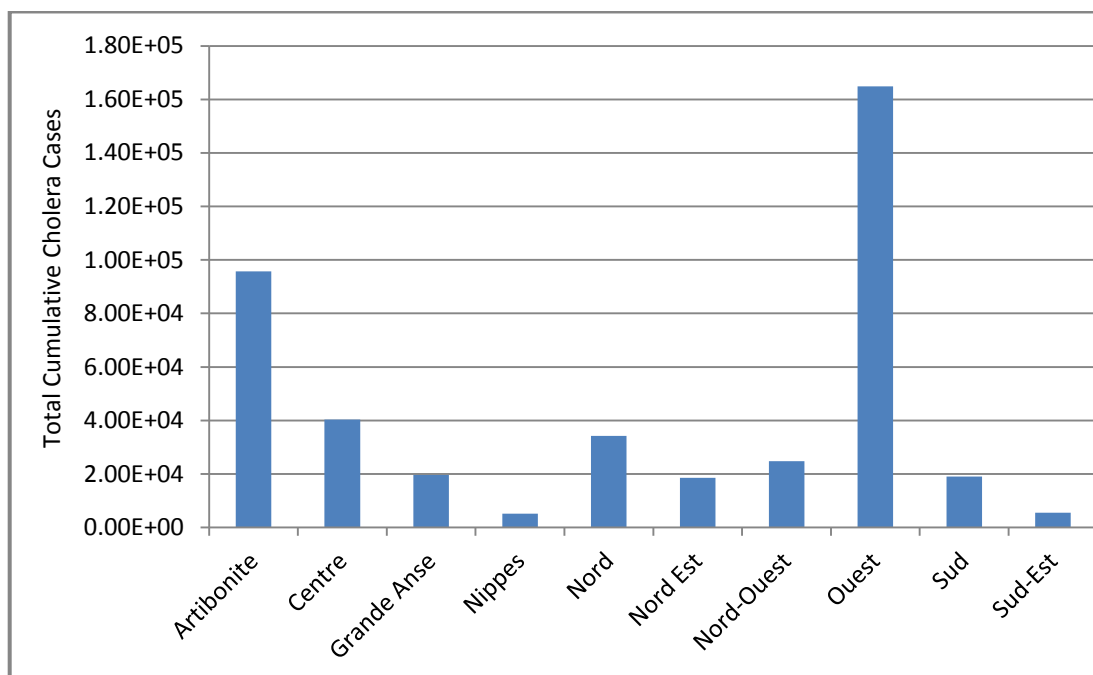
Il primo caso di colera osservato risale al 21 ottobre 2010, vicino a Mirabelais, nel distretto di Centre. Nelle successive cinque settimane l'epidemia si è diffusa in tutti i dipartimenti. Da novembre 2010, si registrano casi anche nella Repubblica Dominicana. Al 22 gennaio 2012, in tutto il territorio di Haiti sono stati registrati 527 012 casi di colera, dei quali 7 025 (1%) hanno portato al decesso [MSPP, 2012]. La veloce diffusione dell'epidemia e il numero di decessi avvenuti nel Paese sono collegati sia a una mancanza di immunità della popolazione a una malattia che ad Haiti non si presentava da più di un secolo, sia a una scarsa preparazione delle strutture sanitarie e del personale medico, che non sono riusciti a contenere il contagio [WHO, 2011].

La variante di colera responsabile dell'epidemia (sierotipo O1, il più letale) ha un decorso molto rapido, soprattutto nei neonati (che possono morire nell'arco di 3-4 ore dal momento del manifestarsi dei sintomi). L'epidemia è iniziata dal fiume Artibonite e progressivamente si è diffusa in tutti i distretti. Si sospetta che il sistema

fluviale sia stato infettato a causa di una fuoriuscita di liquami da un campo nepalese (*Nepalese Annapurna United Nations camp*), facente parte del contingente ONU, che ha smaltito scarichi organici nel fiume Meye, un affluente del fiume Artibonite [*Hendriksen et al.*, 2011].

Secondo il rapporto redatto dal CDC [*CDC*, 2011], al 2 settembre 2011 in tutta Haiti sono stati osservati 433 984 casi, dei quali l'1% ha portato ad un decesso (**Tabella 1.2**). Il dipartimento più colpito è quello di Ovest (**Figura 1.5**). Entrando nel dettaglio, dei 166 344 casi osservati a Ovest, 114 443 sono stati registrati a Port-au-Prince (circa il 69%). Infatti, la maggior parte della popolazione è residente nelle città, o si è trasferita dalle zone rurali dopo il terremoto, per cercare rifugio nei campi di prima accoglienza. Nonostante non sia il distretto in cui si registra il maggior numero di casi, Sud-Est presenta la maggior percentuale di decessi (5%).

A seguito del danneggiamento dei pochi centri sanitari accessibili alla popolazione, le organizzazioni umanitarie gestiscono dei presidi medici nelle tendopoli e negli accampamenti temporanei e distribuiscono kit per l'igiene personale e degli spazi abitativi, per prevenire e arginare il contagio della malattia. Il danneggiamento totale o parziale delle infrastrutture idriche ha provocato in molte zone l'impossibilità per la popolazione di accedere all'acqua potabile, con conseguenti carenze igienico-sanitarie e aumentando il rischio di contagio da colera. Per questo motivo la distribuzione di acqua potabile, la riabilitazione delle infrastrutture idriche di base, la perforazione di nuovi pozzi, la costruzione di latrine, la promozione di programmi di sensibilizzazione sulle buone pratiche di comportamento igienico-sanitario, sono attività prioritarie.



**Figura 1.5:** Grafico dei casi di colera per ogni dipartimento, registrati al 2 settembre 2011 [CDC, 2011].

Dipartimento	Popolazione (ab)	Casi cumulati di colera	Percentuale di popolazione infetta	Decessi cumulati da colera	Percentuale di infetti deceduti
<b>Artibonite</b>	1 361 173	96 943	7%	1 192	1%
<b>Centre</b>	625 235	40 832	7%	521	1%
<b>Grande Anse</b>	368 854	20 544	6%	897	4%
<b>Nippes</b>	270 560	5 301	2%	175	3%
<b>Nord</b>	851 224	34 911	4%	731	2%
<b>Nord Est</b>	337 101	18 853	6%	303	2%
<b>Nord Ouest</b>	482 550	25 067	5%	351	1%
<b>Ouest</b>	3 117 344	166 344	5%	1 514	1%
<b>Sud</b>	508 207	19 348	4%	272	1%
<b>Sud-Est</b>	656 513	5 841	1%	304	5%
<b>Totale</b>	8 578 761	433 984	5%	6 260	1%

**Tabella 1.2:** Per ogni dipartimento sono elencati la popolazione, il numero totale di casi cumulati di colera, la percentuale della popolazione colpita, il totale dei decessi e la percentuale di decessi tra chi ha contratto la malattia. I dati si riferiscono al 2 settembre 2011 [CDC, 2011].

## 1.5 Modelli epidemiologici esistenti

Per poter programmare le misure di intervento in maniera razionale, è utile disporre di modelli quantitativi delle epidemie. Essi, infatti, se opportunamente calibrati sui dati disponibili, forniscono previsioni sull'evoluzione delle malattie sotto diverse ipotesi di controllo.

Il primo modello che ha delineato la dinamica del colera risale al 1979 ed è relativo all'epidemia avvenuta nel Mediterraneo del 1973 [Capasso & Paveri-Fontana, 1979]. Il modello descrive le dinamiche degli individui infetti e del batterio patogeno appartenenti ad una stessa comunità, attraverso un sistema di due equazioni differenziali. Successivamente questo modello è stato esteso includendo anche la dinamica della popolazione suscettibile [Codeço, 2001].

Negli anni sono state proposte diverse modifiche al modello, per esempio, considerando l'iperinfettività dei batteri appena escreti [Hartley et al., 2006], oppure l'interazione preda-predatore tra i batteri *Vibrio cholerae* e i loro fagi [Jensen et al., 2006]. Poiché la propagazione del colera è associata alla presenza del batterio nella riserva d'acqua, è stata poi aggiunta un'ulteriore equazione per descrivere l'evoluzione nel tempo del volume di acqua contaminata [Pascual et al., 2002; Righetto et al., 2012]. Inoltre sono state introdotte anche delle forzanti climatiche e stagionali, per includere l'influenza del differente regime idrico sull'evoluzione dell'epidemia [Colwell, 1996; Lipp et al., 2002; Koelle et al., 2005].

Per descrivere con maggiore accuratezza la diffusione del colera in un territorio, è importante introdurre la componente spaziale, che è collegata alla mobilità umana, alle caratteristiche del territorio e, soprattutto, alla rete idrologica. I modelli che incorporano questa componente sono più recenti. A tale scopo sono state considerate le differenti connessioni idrologiche tra le diverse comunità locali [Bertuzzo et al., 2008; Bertuzzo et al., 2010]. In questi casi, il sistema di equazioni (che descrive la dinamica del compartimento dei suscettibili, degli infetti e dei batteri) è applicato a ogni comunità appartenente all'area di studio e l'equazione della concentrazione batterica è modificata in modo da descrivere il movimento dell'agente patogeno. I fiumi sono i principali vettori di trasmissione di diffusione del batterio *Vibrio*

*cholerae* e infatti, anche ad Haiti, il primo caso è stato osservato vicino al fiume Artibonite.

In molti casi, gli individui infetti sono asintomatici (circa l'80%) e quindi possono inavvertitamente portare il colera da una comunità all'altra. Infatti, la mobilità umana è la causa della propagazione della malattia lungo percorsi alternativi a quelli idrologici. Uno dei metodi per descrivere il ruolo del movimento umano nella diffusione del colera è quello di integrare il modello epidemiologico con un modello gravitazionale [Mari et al., 2011], dove i movimenti tra due comunità locali degli individui dipendono dalla grandezza e dalla distanza delle comunità di appartenenza. Questo modello è già stato testato per altre malattie fortemente influenzate dalla mobilità umana, come l'influenza [Eggo et al., 2010] o la malaria [Xia et al., 2010]. La grande disponibilità, grazie ai *Geographic Information Systems*, di dati georeferenziati su idrologia, rete stradale, distribuzione della popolazione, strutture sanitarie, etc. rende applicabili i modelli spaziali alle diverse situazioni presenti in ogni Paese e per ogni malattia. I modelli SIB (Suscettibili-Infetti-Batteri) spazialmente espliciti integrati con la rete idrologica e la mobilità umana sono stati effettivamente utilizzati per l'epidemia del KwaZulu-Natal, in Sud Africa [Mari et al., 2011] e per quella di Haiti [Bertuzzo et al., 2011]. In quest'ultimo caso, il territorio è stato diviso seguendo i confini amministrativi, considerando unicamente un meccanismo di mobilità umana e non il trasporto su rete idrologica. Solo successivamente si è ricavata la rete idrologica e si è scelta una suddivisione del territorio che ne rispettasse la continuità [Rinaldo et al., 2012].

Più recentemente, per considerare anche il legame tra la stagione delle piogge e la risorgenza dell'epidemia che ha colpito Haiti, nel modello è stata introdotta anche la precipitazione, come forzante esterna che influenza il tasso di contaminazione dell'acqua [Rinaldo et al., 2012].

Lo scopo della tesi è quello di descrivere sulla base di quest'ultimo modello, l'andamento dell'epidemia di colera che ha colpito Haiti ed effettuare delle previsioni sulla sua evoluzione a breve e a lungo termine che siano più accurate di quelle fino ad ora impiegate. A tal fine, si discutono e utilizzano diversi generatori di



precipitazione meteorica, mediante i quali è possibile compiere delle previsioni di precipitazione, che vengono poi inserite come ingresso nel modello epidemiologico.

# 2

## **MODELLO EPIDEMIOLOGICO**

In questo capitolo si descrive il modello utilizzato durante le simulazioni effettuate. Inizialmente si illustra il modello che descrive l'evoluzione dell'epidemia di colera alla scala locale, basandosi sulle interazioni tra quattro compartimenti: i suscettibili, gli infetti, i ristabiliti (e quindi immuni) e i batteri. Successivamente, si introduce un modello spazialmente esplicito, in cui l'area di studio viene divisa in tante unità fondamentali. Per ognuna di esse si sviluppa un modello epidemiologico locale, che viene integrato con le componenti relative al trasporto idrologico e alla mobilità umana che collegano tra loro le unità fondamentali considerando le connessioni spaziali. Inoltre, notando la presenza di una relazione tra gli andamenti della precipitazione e dei casi di colera, si modifica il modello introducendo un termine che descriva possibili meccanismi di recrudescenza dell'epidemia dovuti a fenomeni legati a eventi piovosi.

Infine si illustra il metodo di stima e calibrazione dei parametri presenti nel modello. Infatti, se alcuni di essi sono stati ricavati dalla letteratura esistente, altri sono stati stimati empiricamente, in modo da riprodurre i casi di colera osservati.

## 2.1 Il modello SIRB (Suscettibili-Infetti-Ristabiliti-Batteri)

Si descrive inizialmente il modello epidemiologico locale, in cui successivamente vengono introdotte le dinamiche di trasporto dei batteri e della popolazione.

### 2.1.1 Il modello epidemiologico locale

Il modello epidemiologico locale descrive le dinamiche di avvezione del colera attraverso le interazioni principali che avvengono tra quattro compartimenti:

- *Suscettibili*: sono tutte quelle persone che potrebbero essere infettate. Sono esclusi da questa categoria tutti coloro che sono infetti, o che sono immuni, per aver contratto il colera o per essere stati vaccinati.
- *Infetti/infettivi*: sono coloro che sono affetti dalla malattia e possono trasmetterla. Tra gli infetti circa l'80% non mostrano alcun sintomo [Rinaldo *et al.*, 2012]. Gli asintomatici acquistano una grande importanza nel momento in cui si considera la mobilità umana. Infatti, non essendo debilitati dalla malattia, continuano ad effettuare quegli spostamenti giornalieri che sono una delle cause della diffusione del colera tra le diverse comunità.
- *Ristabiliti*: individui che hanno contratto il colera ma che sono guariti e che di conseguenza hanno sviluppato un'immunità almeno parziale. Dopo un periodo che può variare dalle 12 settimane ai 10 anni [King *et al.*, 2008; Koelle *et al.*, 2005], i ristabiliti perdono l'immunità acquisita, diventano nuovamente suscettibili e possono contrarre nuovamente la malattia.
- *Batteri*: sono i vibrioni nell'ambiente acquatico con cui i suscettibili e gli infetti vengono a contatto attraverso ingestione ed escrezione.

Nella **Figura 2.1** sono rappresentati graficamente i flussi descritti dal modello.

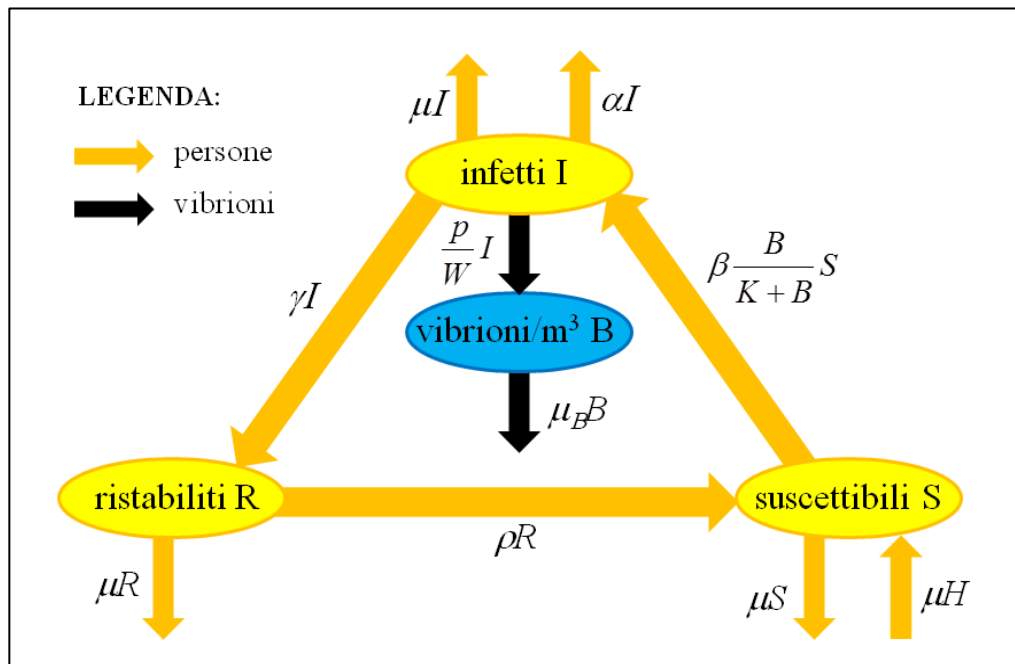


Figura 2.1: Schema del modello epidemiologico locale.

Le variabili e i parametri utilizzati sono:

- S      Abbondanza locale di individui suscettibili
- I      Abbondanza locale di individui infetti
- R      Abbondanza locale di individui ristabiliti
- B      Concentrazione di *Vibrio cholerae*
- H      Totale della popolazione umana all'equilibrio demografico in assenza dell'epidemia
- $\mu$     Tasso di natalità e mortalità demografica della popolazione umana ( $\text{giorni}^{-1}$ )
- $\beta$     Tasso di esposizione all'acqua o al cibo contaminato ( $\text{giorni}^{-1}$ )
- K      Costante di semisaturazione, ovvero la concentrazione di *Vibrio cholerae* nell'acqua per la quale si ha il 50% di possibilità di contrarre il colera ( $\text{vibrioni/m}^3$ )
- $\alpha$     Tasso di mortalità aggiuntivo dovuto al colera ( $\text{giorni}^{-1}$ )
- $\gamma$     Tasso di guarigione ( $\text{giorni}^{-1}$ ), corrispondente all'inverso della durata della malattia

$\mu_B$	Tasso di mortalità di <i>Vibrio cholerae</i> nell'ambiente (giorni <sup>-1</sup> )
$p$	Tasso di contaminazione pro-capite (vibrioni giorni <sup>-1</sup> persona <sup>-1</sup> )
$\rho$	Tasso di perdita dell'immunità (giorni <sup>-1</sup> )
$W$	Volume della riserva d'acqua locale (m <sup>3</sup> )

La formulazione del modello epidemiologico locale è la seguente:

$$\begin{cases} \frac{dS}{dt} = \mu(H - S) - \beta \frac{B}{K + B} S + \rho R \\ \frac{dI}{dt} = \beta \frac{B}{K + B} S - (\gamma + \alpha + \mu) I \\ \frac{dR}{dt} = \gamma I - (\rho + \mu) R \\ \frac{dB}{dt} = -\mu_B B + \frac{p}{W} I \end{cases} \quad (2.1)$$

L'evoluzione del compartimento dei suscettibili (prima equazione) è descritta dal bilancio tra la demografia della popolazione, gli individui contagiati a causa del contatto con l'agente patogeno che contamina la riserva d'acqua e i guariti che perdono l'immunità. La popolazione umana è assunta, in assenza dell'epidemia, all'equilibrio demografico ( $H$ ). Il parametro  $\beta$  rappresenta il tasso di esposizione all'acqua contaminata, mentre  $B/(K + B)$  è una funzione saturante che indica la probabilità di diventare infetti a causa di un'esposizione a una concentrazione di batteri  $B$ . In questo formalismo,  $K$  è la costante di semisaturazione, cioè la concentrazione di patogeni per la quale la probabilità di un suscettibile di diventare infetto è la metà del valore massimo.

La dinamica del compartimento degli infetti (seconda equazione) è data dal bilancio tra il flusso entrante di nuovi malati e quello uscente dovuto alle guarigioni, alla mortalità naturale, considerando anche un tasso di mortalità aggiuntivo dovuto al colera ( $\alpha$ ).

La terza equazione, che descrive la dinamica dei ristabiliti (ovvero tutti coloro che sono guariti, sono immuni e non possono contrarre il colera), è data dal bilancio tra i guariti dal colera e coloro che muoiono per cause naturali o che perdono l'immunità.

Il parametro  $\rho$  quantifica il tasso col quale l'immunità viene persa e, quindi, la relativa ripopolazione del serbatoio dei suscettibili.

L'ultima equazione descrive la concentrazione di batteri presenti nella riserva d'acqua locale. Gli infetti contribuiscono ad aumentare la concentrazione di vibroni con un tasso  $p/W$ , dove  $p$  è il tasso di batteri escreti da un individuo infetto che raggiunge e contamina la riserva d'acqua locale di volume  $W$ . Si assume che il tasso di crescita dei batteri nell'ambiente naturale ( $\mu_B$ ) sia negativo, ossia che, in assenza di escrezione da parte di individui infetti, i batteri tendano a estinguersi.

Le condizioni iniziali sono  $X(0) = [S_0, I_0, R_0, B_0]^T$ , dove,  $S_0 = H - I_0$ ,  $I_0$  è il numero di infetti iniziali, stabiliti sulla base delle osservazioni in campo,  $R_0 = 0$  e  $B_0$  è la concentrazione di batteri iniziale, calcolata come la concentrazione di quasi equilibrio  $B_0 = \frac{pI_0}{W\mu_B}$ .

Sotto certe condizioni, il modello presenta, oltre all'equilibrio corrispondente alla popolazione sana alla capacità portante  $\bar{X}_H = [H, 0, 0, 0]^T$ , un secondo equilibrio corrispondente alla popolazione parzialmente infetta  $\bar{X}_+ = [\bar{S}_+, \bar{I}_+, \bar{R}_+, \bar{B}_+]^T$ . Quest'ultimo equilibrio ha però senso biologico solo se  $\bar{I}_+$  non è negativo, quindi deve essere verificata la condizione  $R_0 \geq 1$ , dove  $R_0$  è il tasso base di riproduzione della malattia, che può essere interpretato come il numero medio di infezioni secondarie prodotte da un individuo infetto immesso in una popolazione sana e costituita da  $H$  individui ed è definito come:

$$R_0 = \frac{\beta p H}{(\gamma + \alpha + \mu) K \mu_B W}$$

Se  $R_0$  è minore di 1, l'equilibrio con la popolazione parzialmente infetta non può sussistere. Si può dimostrare che in quest'ultimo caso l'equilibrio  $\bar{X}_H$  è stabile qualsiasi siano le condizioni iniziali e quindi l'epidemia va estinguendosi naturalmente. Invece, se  $R_0$  è maggiore di 1, l'equilibrio corrispondente alla popolazione parzialmente infetta  $\bar{X}_+$  è l'unico equilibrio stabile e quindi la malattia è endemica, ovvero una frazione costante della popolazione è infetta [Codeço, 2001].

## 2.1.2 Il modello spazialmente esplicito

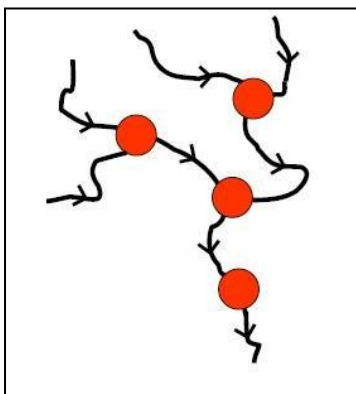
Quando si studiano situazioni a grande scala, come nel caso di Haiti, un modello locale non descrive adeguatamente le dinamiche spazio temporali. È necessario dividere il territorio in unità fondamentali nell'ambito delle quali il modello locale sia almeno approssimativamente corretto e poi bisogna connettere tali unità. In particolare, per il caso di studio, sono state scelte delle entità idrologiche (**Capitolo 3.3**). Per descrivere i legami tra le diverse entità si utilizza un grafo, i cui nodi sono le entità stesse e gli archi rappresentano le loro connessioni (**Figura 2.2**). Ad ogni nodo corrisponde un modello epidemiologico locale (**Modello 2.1**), a cui vanno aggiunti termini di trasporto e di mobilità che collegano tale nodo a quelli ad esso connessi. I parametri di tale modello sono assunti omogeneamente distribuiti nello spazio. Infatti, il vantaggio di questo modello è anche quello di riuscire a descrivere le dinamiche su larga scala con un grande livello di dettaglio ma con un numero limitato di gradi di libertà, comportando quindi in aggiunta una diminuzione significativa dei costi computazionali.



**Figura 2.2:** Rappresentazione grafica del modello spazialmente esplicito, che rappresenta il territorio per mezzo di nodi e archi.

Si assume che il batterio *Vibrio cholerae* si diffonda essenzialmente per mezzo della rete idrologica e tramite i movimenti degli individui tra le diverse comunità. Infatti, il *Vibrio cholerae* può sopravvivere nell'ambiente acquatico anche in associazione con lo zooplancton o con la vegetazione acquatica [Colwell, 1996] e quindi si propaga lungo la rete idrologica anche insieme a questi organismi. La mobilità umana è invece basata sui movimenti giornalieri della popolazione. I meccanismi di contaminazione associati al movimento umano sono due: da una parte gli individui suscettibili appartenenti a nodi non infetti possono contrarre il colera durante lo spostamento in altre comunità e, una volta tornati nel proprio nodo di origine, possono svolgere il ruolo di agenti contaminanti della propria riserva idrica; dall'altra, gli individui infetti (generalmente asintomatici) possono diffondere il batterio durante i loro spostamenti quotidiani in comunità differenti dalla propria. L'alto rapporto tra asintomatici e sintomatici rende ragionevole ipotizzare che il movimento delle persone affette da colera non sia sempre impedito.

#### Trasporto idrologico



**Figura 2.3:** Modellizzazione del trasporto idrologico. In rosso sono rappresentati i nodi della rete, mentre in nero le connessioni idrologiche.

Il territorio di Haiti è suddiviso in unità idrologiche, ognuna delle quali costituisce un nodo del grafo su cui si basa il modello spazialmente esplicito. Ogni nodo è connesso agli altri coerentemente con la rete idrologica (**Figura 2.3**). La diffusione di *Vibrio cholerae* lungo la rete idrologica è descritta come un cammino casuale su un grafo orientato. I batteri hanno un tasso di mobilità  $l$  (espresso in  $\text{giorni}^{-1}$ ), dove  $1/l$  è il tempo di residenza medio dei batteri in ciascuna riserva d'acqua. I batteri si muovono dal nodo  $i$  al nodo ad esso interconnesso  $j$  con una probabilità pari a:



$$P_{ij} = \begin{cases} \frac{P_{out}}{d_{out}(i)P_{out} + d_{in}(i)P_{in}} & \text{se } i \rightarrow j \\ \frac{P_{in}}{d_{out}(i)P_{out} + d_{in}(i)P_{in}} & \text{se } i \leftarrow j \\ 0 & \text{se } i \leftrightarrow j \end{cases}$$

dove:

$P_{out}$  è la probabilità di *Vibrio cholerae* di uscire dal nodo  $i$

$P_{in}$  è la probabilità di *Vibrio cholerae* di entrare nel nodo  $i$

$d_{out}$  è il numero di archi uscenti

$d_{in}$  è il numero di archi entranti

Il processo di trasporto è conservativo, quindi:

$$\sum_{j=1}^{d(i)} P_{ij} = 1$$

dove  $P_{ij}$  è la probabilità che i batteri vadano dal nodo  $i$  al nodo  $j$ , mentre  $d(i)$  è dato dalla somma degli archi entranti e di quelli uscenti nel nodo  $i$  ( $d(i) = d_{out}(i) + d_{in}(i)$ ). Inoltre si può individuare un errore (o *bias*) di trasporto come:

$$b = P_{out} - P_{in} = 2P_{out} - 1$$

Nel caso di studio si assume che il trasporto avvenga solo da monte verso valle, in quanto il trasporto da valle verso monte è trascurabile ( $P_{in} = 0$  e  $P_{out} = 1$ ).

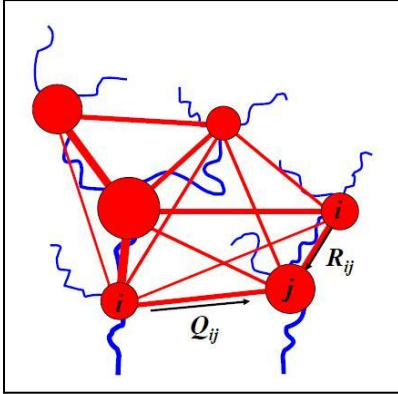
Di conseguenza il modello epidemiologico locale (**Modello 2.1**) assume la seguente forma:

$$\begin{cases} \frac{dS_i}{dt} = \mu(H_i - S_i) - \beta \frac{B_i}{K + B_i} S_i + \rho R_i \\ \frac{dI_i}{dt} = \beta \frac{B_i}{K + B_i} S_i - (\gamma + \alpha + \mu) I_i \\ \frac{dR_i}{dt} = \gamma I_i - (\rho + \mu) R_i \\ \frac{dB_i}{dt} = -\mu_B B_i + \frac{p}{W_i} I_i - l \left( B_i + \sum_{j=1}^n P_{ji} B_j \frac{W_j}{W_i} \right) \end{cases} \quad (2.2)$$

Le condizioni iniziali dei nodi infetti sono  $X_i(0) = [S_{i0}, I_{i0}, 0, B_{i0}]^T$ , con  $S_{i0} = H_i - I_{i0}$ , dove  $I_{i0}$  è il numero di infetti iniziali, osservati nella realtà, e  $B_{i0}$  è la concentrazione di batteri iniziale, calcolata come la concentrazione di quasi equilibrio  $B_{i0} = \frac{pI_{i0}}{W_i\mu_B}$ . Invece, le condizioni iniziali dei nodi non contagiati sono  $X_i(0) = [H_i, 0, 0, 0]^T$ , poiché non sono presenti individui infetti. Il **Modello 2.2** presuppone la conoscenza della popolazione presente in ogni nodo. Per i nodi al di fuori del territorio di Haiti è impostato di *default* il valore 1, in modo da poter soddisfare il sistema soprastante. Il volume della riserva d'acqua locale è considerato essere proporzionale alla popolazione  $W_i = cH_i$ , in quanto si presuppone che una città con più abitanti abbia a disposizione riserve d'acqua maggiori [Mari *et al.*, 2011]. Si ottiene quindi un sistema di  $4 \cdot n$  equazioni, dove  $n$  è il numero dei nodi. Si noti che ponendo  $l = 0$ , il modello si riduce a un insieme di equazioni locali disgiunte.

### *Mobilità umana*

La rete di mobilità umana è rappresentata da un grafo nel quale i nodi corrispondono a quelli della rete idrologica, mentre gli archi rappresentano tutte le possibili connessioni tra un nodo e tutti gli altri nodi che compongono la rete (**Figura 2.4**). A causa della conformazione di Haiti, in questo contesto non è realistico considerare la distanza euclidea tra due punti generici (per esempio per nodi separati dal mare). Per questa ragione, è stato prima creato un grafo nel quale ogni nodo è collegato soltanto con i quattro più vicini. Le distanze su questo grafo sono euclidee. La distanza tra due nodi generici  $i$  e  $j$  è calcolata come la lunghezza del cammino più breve, lungo il grafo precedente, che connette  $i$  e  $j$ . Il cammino più breve è stato



**Figura 2.4:** Modellizzazione del trasporto dovuto alla mobilità umana (archi rossi), sovrapposto alla rete componente il trasporto idrologico (in blu).

trovato con l'algoritmo standard Dijkstra. La rete di mobilità umana è dunque definita attraverso un grafo completamente connesso.

Il flusso di mobilità umana tra due nodi è descritto con un modello gravitativo, in cui gli individui lasciano il loro nodo di origine  $i$  con una probabilità  $m$ , e raggiungono il nodo di destinazione  $j$  con una probabilità  $Q_{ij}$  e poi tornano al proprio nodo di partenza. In analogia col modello gravitazionale, la probabilità  $Q_{ij}$  è definita come:

$$Q_{ij} = \frac{H_j e^{-\frac{d_{ij}}{D}}}{\sum_{k \neq i}^n H_k e^{-\frac{d_{ik}}{D}}}$$

dove  $H_j$  e  $H_k$  sono le popolazioni rispettivamente del nodo  $j$  e del nodo  $k$ ,  $d_{ij}$  e  $d_{ik}$  sono le distanze tra i nodi  $i$  e  $j$  e tra i nodi  $i$  e  $k$ . Infine  $D$  è la distanza di dispersione media e  $n$  il numero di nodi. In questo modello il nodo  $j$  è tanto più attrattivo per il nodo  $i$  quanto più è grande la popolazione di  $j$  e tanto più è piccola la distanza che li separa ( $d_{ij}$ ). La capacità di movimento è differente da nazione a nazione e dipende dalle infrastrutture, dalla topografia e dalle abitudini locali, quindi il parametro  $D$  deve essere calibrato a seconda della situazione. Poiché  $Q_{ij}$  è una probabilità, deve essere:

$$\sum_{j=1}^n Q_{ij} = 1$$

Tenendo presente le considerazioni fatte finora, il modello SIRB comprendente sia il trasporto idrologico che la mobilità umana, risulta essere il seguente:

$$\begin{cases} \frac{dS_i}{dt} = \mu(H_i - S_i) - \beta \left[ (1-m) \frac{B_i}{K+B_i} + m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} \right] S_i + \rho R_i \\ \frac{dI_i}{dt} = \beta \left[ (1-m) \frac{B_i}{K+B_i} + m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} \right] S_i - (\gamma + \alpha + \mu) I_i \\ \frac{dR_i}{dt} = \gamma I_i - (\rho + \mu) R_i \\ \frac{dB_i}{dt} = \frac{p}{W_i} \left[ (1-m) I_i + m \sum_{j=1}^n Q_{ji} I_j \right] - \mu_B B_i - l \left( B_i - \sum_{j=1}^n P_{ji} \frac{W_j}{W_i} B_j \right) \end{cases} \quad (2.3)$$

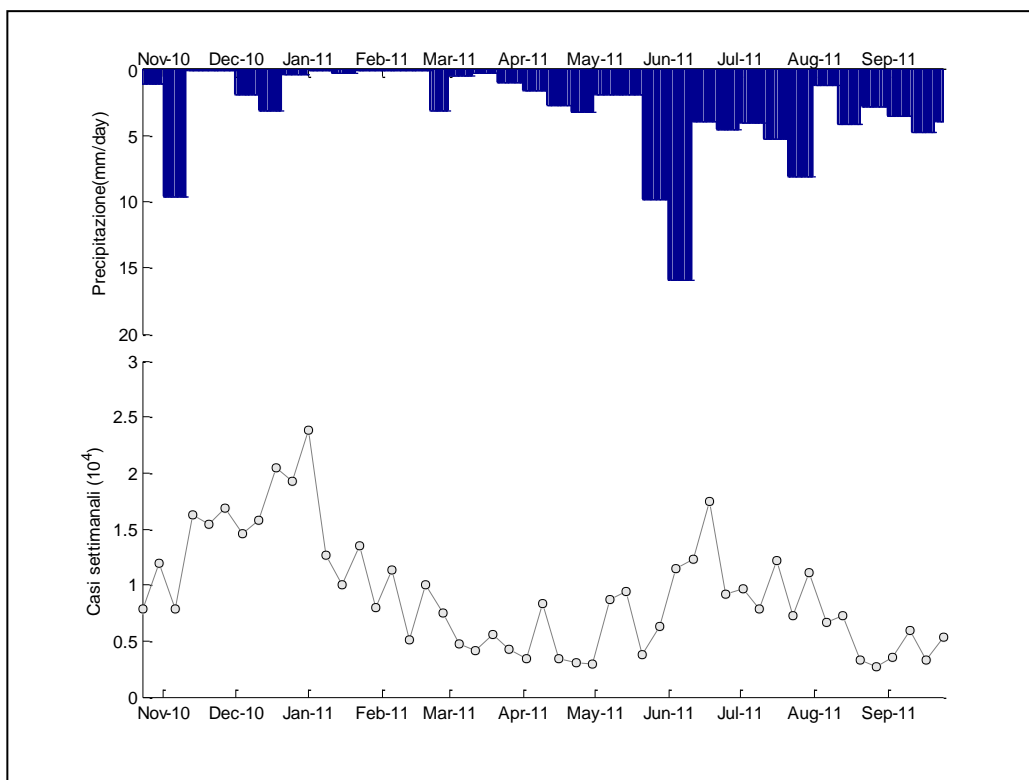
dove  $m$  è la parte di popolazione di un nodo che viaggia quotidianamente e viceversa,  $1-m$  è la parte di popolazione che rimane stanziale nel nodo stesso. Per quanto riguarda la prima equazione, il termine  $\beta m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} S_i$  è la frazione di suscettibili che contrae la malattia viaggiando. Il termine  $\frac{p}{W_i} m \sum_{j=1}^n Q_{ji} I_j$ , appartenente alla quarta equazione, rappresenta la quantità di batteri introdotta nel nodo  $i$  da tutti gli infetti asintomatici appartenenti ad altri nodi che viaggiano quotidianamente verso il nodo  $i$ .

## 2.2 Il modello SIRB con le precipitazioni

Nel grafico in **Figura 2.5** sono riportati i casi di colera registrati settimanalmente e la precipitazione decadale nel periodo corrispondente. I casi di colera presentano due picchi successivi. Infatti, ad un primo *exploit* iniziale di infetti, ne è seguito un secondo dopo qualche mese. Il primo picco di casi di colera è più alto, in quanto ad Haiti non si riportavano casi di colera da più di un secolo e quindi l'intera popolazione risultava essere suscettibile. Col tempo i casi settimanali registrati di infetti diminuiscono (periodo che va da febbraio ad aprile 2011) poiché chi ha contratto la malattia diventa, una volta guarito, immune, e quindi il numero di suscettibili diminuisce. Ciononostante, si registra un secondo incremento dei casi di colera nel periodo tra maggio e agosto 2011.

Il **Modello 2.3** non è in grado di riprodurre questa particolare evoluzione dell'epidemia, neanche assumendo un tasso di perdita dell'immunità  $\rho$  maggiore e,

quindi, un più veloce ripopolamento del serbatoio dei suscettibili. Di conseguenza, l'insorgenza di un secondo picco va attribuita ad una forzante esterna. Osservando la **Figura 2.5** è plausibile presumere che sia presente una relazione tra le precipitazioni stagionali e la risorgenza della malattia, infatti il secondo incremento dei casi di colera avviene proprio in corrispondenza dell'inizio del periodo piovoso. In seguito a queste considerazioni, si introduce nel modello una descrizione dell'aumento del tasso di contaminazione dell'acqua da *Vibrio cholerae* causato dall'incremento dell'intensità delle precipitazioni.



**Figura 2.5:** Nel grafico in alto sono riportate le precipitazioni nel periodo dal 23 ottobre 2010 fino al 30 settembre 2011. Nel grafico in basso sono rappresentati i casi di colera settimanali registrati. È possibile notare una corrispondenza tra l'inizio del periodo piovoso a maggio/giugno e il secondo picco di casi di colera.

Il modello che tiene in considerazione le piogge come forzante esterna risulta essere il seguente:

$$\begin{cases} \frac{dS_i}{dt} = \mu(H_i - S_i) - \beta \left[ (1-m) \frac{B_i}{K+B_i} + m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} \right] S_i + \rho R_i \\ \frac{dI_i}{dt} = \beta \left[ (1-m) \frac{B_i}{K+B_i} + m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} \right] S_i - (\gamma + \alpha + \mu) I_i \\ \frac{dR_i}{dt} = \gamma I_i - (\rho + \mu) R_i \\ \frac{dB_i}{dt} = -\mu_B B_i - l \left( B_i - \sum_{j=1}^n P_{ji} \frac{W_j}{W_i} B_j \right) + \frac{p}{W_i} [1 + \lambda J_i(t)] [(1-m) I_i + m \sum_{j=1}^n Q_{ji} I_j] \end{cases} \quad (2.4)$$

Per giustificare la stagionale risorgenza dell'epidemia di colera, si assume che il tasso di contaminazione dell'acqua  $p$  possa aumentare con la precipitazione  $J_i(t)$ , attraverso un coefficiente  $\lambda$ . Questo è dovuto al fatto che la precipitazione causa un aumento dell'agente patogeno nella riserva d'acqua, poiché attraverso il deflusso superficiale della precipitazione vengono mobilitati grandi quantità di batteri dai siti di defecazione all'aria aperta verso corpi d'acqua utilizzati come riserva potabile e per uso domestico [Rinaldo et al., 2012].

### 2.3 Stima e calibrazione dei parametri

Il modello descritto simula l'andamento dell'epidemia, considerando come infetti sia gli individui sintomatici che quelli asintomatici, ma in realtà i dati disponibili riguardano soltanto i casi osservati (che sono quelli sintomatici). Per poter confrontare le simulazioni effettuate con i dati reali, si definisce il numero di casi cumulati simulati sintomatici per ogni nodo:

$$\frac{dC_i}{dt} = \sigma \cdot \beta \left[ (1-m) \frac{B_i}{K+B_i} + m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} \right] S_i \quad (2.5)$$

Il fattore  $\sigma$  corrisponde alla frazione di individui infetti che, presentando i sintomi del colera, si rivolgono a strutture ospedaliere (circa il 20% degli infetti). Il secondo fattore rappresenta il flusso di nuovi individui infetti, calcolati sommando tutte le

persone che non viaggiano e che quindi sono state contagiate a causa dei batteri presenti nella propria comunità, e tutte quelle che sono state infettate durante il loro viaggio giornaliero (e quindi da batteri presenti in altri nodi). Nel raffronto tra dati reali e simulati si utilizzano i casi cumulati, in quanto questi ultimi sono più stabili rispetto a quelli settimanali, facilitando così le stime dei parametri. Alcuni dei parametri del **Modello 2.4** sono stati ricavati dalla letteratura esistente, mentre altri sono stati stimati tramite calibrazione.

I parametri derivati dalla letteratura esistente sono riportati in **Tabella 2.1**. Il tasso di mortalità/natalità della popolazione  $\mu$  è definito come l'inverso dell'aspettativa di vita media, che, secondo la *CIA* (2009), per un haitiano risulta essere pari a 61 anni e 5 mesi circa. Si assume che le persone possano essere esposte all'acqua o al cibo contaminato ( $\beta$ ) almeno una volta al giorno, così come anche suggerito da *Codeço* (2001), *Bertuzzo et al.* (2008), *Mari et al.* (2011) e *Tuite et al.* (2011). Inoltre, secondo le prime tre fonti, *Bertuzzo et al.* (2011) e *Andrews & Basu* (2011), la durata media della malattia risulta essere di 5 giorni, da cui, calcolando l'inverso, si può ricavare il tasso di guarigione  $\gamma$ . Sempre dagli stessi autori, si desume che la

Parametro	Valore	Unità di misura	Referenze
$\mu$	$1/(61.4 \cdot 365)$	giorni <sup>-1</sup>	[1]
$\beta$	1.0	giorni <sup>-1</sup>	[2], [3], [4], [5]
$\rho$	$1/(3 \cdot 365)$	giorni <sup>-1</sup>	[6]
$\gamma$	0.20	giorni <sup>-1</sup>	[2], [3], [7], [8], [4]
$\alpha$	$4.0 \cdot 10^{-3}$	giorni <sup>-1</sup>	[9]
$\mu_B$	0.2	giorni <sup>-1</sup>	[2], [3], [4], [7]
$\sigma$	0.2	-	[10], [11], [8], [12]

[1] = [ <i>CIA</i> , 2009]	[7] = [ <i>Bertuzzo et al.</i> , 2011]
[2] = [ <i>Codeço</i> , 2001]	[8] = [ <i>Andrews &amp; Basu</i> , 2011]
[3] = [ <i>Bertuzzo et al.</i> , 2008]	[9] = [ <i>PAHO</i> , 2011]
[4] = [ <i>Mari et al.</i> , 2011]	[10] = [ <i>WHO</i> , 2010]
[5] = [ <i>Tuite et al.</i> , 2011]	[11] = [ <i>Longini et al.</i> , 2007]
[6] = [ <i>Koelle et al.</i> , 2005]	[12] = [ <i>Chao et al.</i> , 2011]

**Tabella 2.1:** Valore dei parametri stimati dalla letteratura esistente.

mortalità dei batteri di *Vibrio cholerae* ( $\mu_B$ ) nell'ambiente naturale risulta essere di 0.2 giorni<sup>-1</sup>. Il tasso di mortalità aggiuntivo dovuto al colera ( $\alpha$ ) è fornito dalla PAHO (2011) e risulta essere di 0.004 giorni<sup>-1</sup>. Infine, in accordo con Koelle et al. (2005), si stabilisce che la perdita di immunità  $\rho$  avviene dopo 3 anni.

I parametri rimanenti non possono essere dedotti dalla letteratura esistente, ma devono essere opportunamente stimati. La calibrazione è stata effettuata dal *Laboratory of Ecohydrology* dell'*Ecole Polytechnique Fédérale de Lausanne* (EPFL), che ha utilizzato un approccio basato sul metodo *Markov Chain Monte Carlo* (MCMC). I metodi MCMC sono una famiglia di algoritmi per l'esplorazione della funzione di densità di probabilità a posteriori di una distribuzione di probabilità desiderata, che quindi per il caso di studio è quella del *set* di parametri da calibrare. Per ogni parametro da calibrare si costruisce una catena di *Markov* che esplora diversi valori del parametro e, dopo un certo numero di passi, tende a un equilibrio. Generalmente il problema maggiore è decidere quanti passi sono necessari per convergere a una distribuzione stazionaria con un errore accettabile. Questo metodo permette l'esplorazione di uno spazio di parametri, in cui i campioni casuali esaminati sono correlati. La catena di *Markov* è costruita in modo da avere, quando raggiunge l'equilibrio, il valore del parametro desiderato. Nelle simulazioni è stato usato l'algoritmo DREAM (*Differential Evolution Adaptive Metropolis*), che implementa diverse catene contemporaneamente per assicurare un'intera esplorazione dello spazio dei parametri. Più nello specifico, è stato utilizzato l'algoritmo DREAM<sub>ZS</sub>, una variante del DREAM, che sfrutta i campioni dei passi precedenti delle catene di *Markov* e riduce il numero effettivo di catene parallele necessarie per l'esplorazione dello spazio dei parametri. Per ogni parametro, l'algoritmo, che è inizializzato a priori con distribuzioni uniformi, procede finché non arriva alla convergenza delle catene.

Per la calibrazione sono stati utilizzati i dati dei casi cumulati di colera e delle precipitazioni decadali (fornite dalla *NASA-JAXA's Tropical Rainfall Measuring Mission*) registrati dal 23 ottobre 2010 fino al 28 maggio 2011. Vengono calcolate le catene di *Markov* relative ai parametri da calibrare, che sono il tasso base di



riproduzione della malattia ( $R_0$ ), il tasso di dispersione dei batteri *Vibrio cholerae* ( $l$ ), la frazione di popolazione di un nodo che si sposta quotidianamente in altri nodi ( $m$ ) e il coefficiente  $\lambda$ . In particolare, per ogni parametro sono state implementate tre catene di *Markov*, che dopo 6 590 passi sono arrivate a convergenza. I valori dei parametri calibrati, ossia quelli che garantiscono il minor valore dello scarto quadratico tra casi cumulati osservati e simulati, sono riportati in **Tabella 2.2**.

Parametro	Valore	Unità di misura
$R_0$	4.9174	-
$l$	6.5501	giorni <sup>-1</sup>
$m$	0.6761	-
$\lambda$	2.5154	giorni /mm

**Tabella 2.2:** Valore dei parametri calibrati.

Il tasso aggregato di contaminazione ( $\theta$ ) è derivabile da  $R_0$  come  $\theta = \frac{p}{KW} = \frac{R_0(\gamma+\alpha+\mu)\mu_B}{\beta} = 0.2007$  giorni<sup>-1</sup>. Si riporta di seguito una tabella riassuntiva di tutti i parametri utilizzati nel modello (**Tabella 2.3**).

Parametro	Valore	Unità di misura	Tipo
$\mu$	1/(61.4 · 365)	giorni <sup>-1</sup>	Da letteratura
$\beta$	1.0	giorni <sup>-1</sup>	Da letteratura
$\rho$	1/(3 · 365)	giorni <sup>-1</sup>	Da letteratura
$\gamma$	0.20	giorni <sup>-1</sup>	Da letteratura
$\alpha$	4.0 · 10 <sup>-3</sup>	giorni <sup>-1</sup>	Da letteratura
$\mu_B$	0.2	giorni <sup>-1</sup>	Da letteratura
$\sigma$	0.2	-	Da letteratura
$R_0$	4.9174	-	Calibrato
$l$	6.5501	giorni <sup>-1</sup>	Calibrato
$m$	0.6761	-	Calibrato
$\lambda$	2.5154	giorni /mm	Calibrato
$D$	100	km	Calibrato
$\theta$	0.2007	giorni <sup>-1</sup>	Calibrato

**Tabella 2.3:** Riepilogo dei parametri utilizzati.

# 3

## DATI

L'applicazione dei modelli descritti nel capitolo precedente al caso dell'epidemia di Haiti prevede l'utilizzo di una base di dati composta da:

- i casi di colera registrati, in modo da poter confrontare i risultati delle simulazioni effettuate con quanto effettivamente osservato nel caso in analisi
- la distribuzione della popolazione
- le mappe digitali di elevazione del terreno (DEM) che vengono utilizzate per l'estrazione della rete idrologica
- i dati relativi alle precipitazioni, necessari per l'implementazione del modello di simulazione dell'epidemia di colera (**Capitolo 2, Modello 2.4**) e per la parametrizzazione dei modelli di previsione di precipitazione (**Capitolo 4**); inoltre sono indispensabili per realizzare delle statistiche sull'andamento stagionale delle precipitazioni stesse.

### 3.1 Dati epidemiologici

I dati epidemiologici sono stati forniti dal *Ministère de la Santé Publique et de la Population* (MSPP) di Haiti e dalla *Pan American Health Organization* (PAHO). I dati consistono in rapporti settimanali dei casi di colera osservati per ogni distretto, in cui si riportano i casi registrati dagli enti sanitari. Di conseguenza, i rapporti non comprendono gli infetti asintomatici, poiché non necessitano di cure. Si ricorda che generalmente l'80% degli individui infettati non sviluppa sintomi di alcun genere.

### 3.2 Dati relativi alla popolazione

I dati relativi alla popolazione sono in formato raster e sono stati forniti dall'*Oak Ridge National Laboratory* (ORNL)<sup>6</sup>, che ha utilizzato un approccio innovativo che unisce le potenzialità dei *Geographic Information Systems* con quelle del *Remote Sensing*, per calcolare la distribuzione della popolazione. Questo metodo, implementato dall'ORNL nell'algoritmo *LandScan*, distribuisce i dati del censimento, che sono raccolti secondo i confini amministrativi, sul territorio, considerando la copertura del suolo dell'area di interesse. In particolare, nel momento in cui si suddivide la popolazione, ad ogni copertura viene attribuito un peso diverso (ad esempio alle aree urbanizzate si assegna un peso maggiore rispetto alle aree agricole).

I dati hanno approssimativamente 1 km<sup>2</sup> di risoluzione (30 x 30 arcosecondi, ovvero 0.0083 x 0.0083 gradi) e rappresentano la distribuzione della popolazione mediata su 24 ore. Quest'ultima misura è ottenuta rilevando i dati di distribuzione della popolazione per ogni ora della giornata ed effettuandone la media. In questo

---

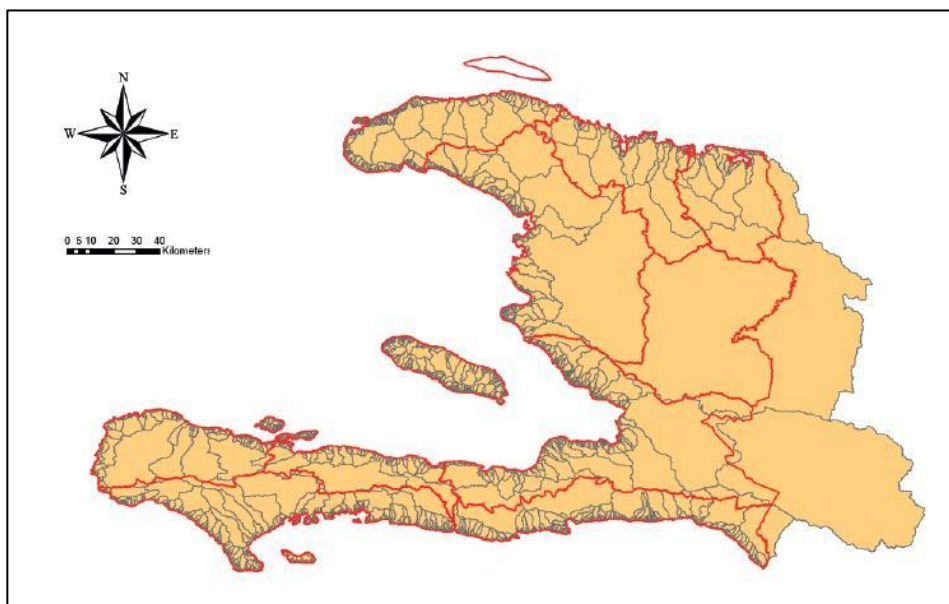
<sup>6</sup> L'*Oak Ridge National Laboratory* (ORNL) è un laboratorio nazionale di scienze interdisciplinari e tecnologia, nato dalla collaborazione tra l'Università del Tennessee e il *Battelle Memorial Institute* e gestito dal Dipartimento dell'Energia degli Stati Uniti, di cui rappresenta il più grande laboratorio di scienza ed energia.

modo si integrano i movimenti diurni e le abitudini di viaggio in una sola misura [ORNL].

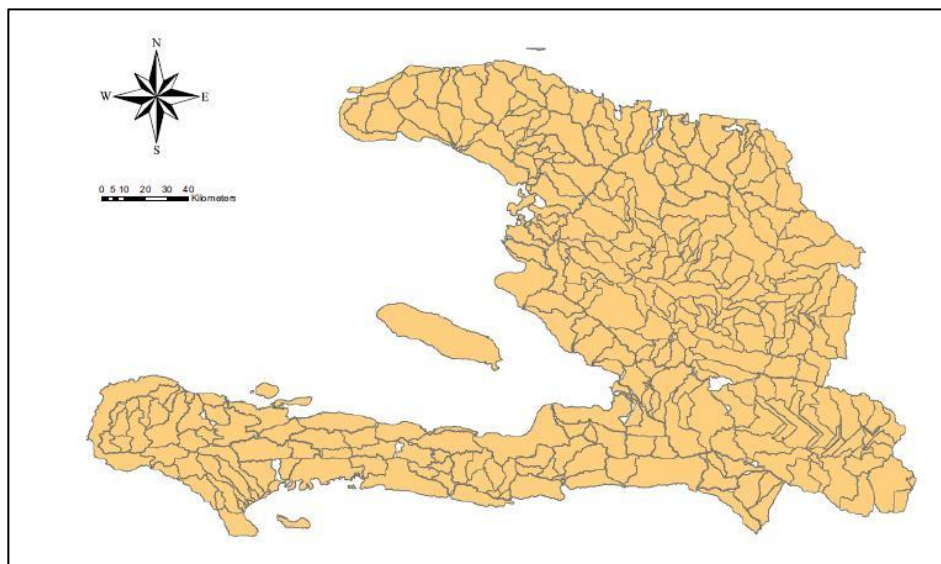
### 3.3 Dati utilizzati per l'estrazione della rete idrologica

Il dato di partenza che permette l'estrazione della rete idrologica è rappresentato dalle mappe digitali di elevazione del terreno (*Digital Terrain Model*, DEM), forniti da *United States Geological Survey* (USGS). I dati disponibili si riferiscono all'area corrispondente all'intera isola di Hispaniola e sono di tipo raster, quindi sono visualizzabili come una griglia georeferenziata in cui ogni cella contiene il relativo dato di elevazione. La precisione della griglia è di 100 metri. Il sistema spaziale di riferimento è GCS WGS 84.

Per questo lavoro è stata considerata l'area che impatta idrologicamente su Haiti. Pertanto, per preservare la continuità idrologica dei corsi d'acqua, è stato necessario includere anche parte della Repubblica Dominicana. L'obiettivo è quello di suddividere il territorio in sottobacini idrologici, ognuno dei quali corrisponde a un nodo del modello. Il territorio è stato frazionato attraverso l'uso di TauDEM, un'estensione del *software* ArcGIS®, che consente di fare un'analisi idrologica a partire da informazioni topografiche. A causa della disomogeneità topografica, il territorio è stato diviso in 9 790 bacini con un'area compresa tra 1 km<sup>2</sup> e 5 000 km<sup>2</sup>, corrispondenti al bacino di Artibonite (**Figura 3.1**). La situazione ideale sarebbe quella in cui il territorio viene diviso in 300 o 400 bacini di area paragonabile. Poiché l'area totale di Haiti risulta essere di circa 27 750 km<sup>2</sup>, l'area media ideale dei bacini dovrebbe essere compresa tra i 69 e i 93 km<sup>2</sup>. Pertanto, è necessario aggregare tra loro i bacini più piccoli o disaggregare quelli più grandi. Il risultato di questo processo è riportato nella **Figura 3.2** dove sono visualizzati i 366 bacini finali, ognuno dei quali costituisce un nodo della rete [Biermann, 2011]. A ogni nodo sono poi assegnate le coordinate del centro di massa della distribuzione della popolazione del corrispondente bacino. Infine, per definire gli archi corrispondenti ai flussi di mobilità, vengono calcolate le distanze tra i nodi (**Capitolo 2.1.2**).



**Figura 3.1:** Mappa dei bacini idrologici in cui è stato inizialmente suddiviso il territorio di Haiti. In arancio è indicata l'area che impatta idrologicamente su Haiti, in rosso sono delineati i confini amministrativi dei dieci distretti, mentre in grigio sono rappresentati i confini dei bacini.



**Figura 3.2:** Mappa relativa ai 366 bacini finali in cui è stata divisa Haiti. È possibile notare l'omogeneità delle aree dei bacini stessi e della loro distribuzione. In arancio è indicata l'area che impatta idrologicamente su Haiti, in grigio sono rappresentati i confini dei bacini.

### 3.4 Precipitazioni

I dati di precipitazione a disposizione sono in formato sia decadale che giornaliero. Per quanto riguarda i dati decadali, ogni mese è suddiviso in tre decadi, di cui l'ultima è di lunghezza variabile. I dati decadali sono stati forniti da due diverse fonti:

- dalla *NASA-JAXA's Tropical Rainfall Measuring Mission*<sup>7</sup> attraverso il geoportale FEWS NET<sup>8</sup>, messo a disposizione da USGS. Questi dati coprono il periodo dall'1 giugno 2010 fino al 30 settembre 2011.
- dal geoportale <http://portal.iri.columbia.edu/portal/server.pt> dell'*International Research Institute for Climate and Society*, un ente derivato dalla cooperazione tra il *NOAA's Climate Program Office* e la *Columbia University*. Da questa fonte è possibile reperire dieci anni di dati decadali di precipitazione, dal 1 gennaio 2002 al 31 luglio 2011 (**Figura 3.3**).

In entrambi i casi, i dati utilizzati sono spazialmente distribuiti, con una risoluzione di 0.25 gradi di latitudine e longitudine. Ogni immagine è composta da 12x27 celle, cosicché l'area totale coperta è circa pari a 330x750 km<sup>2</sup>.

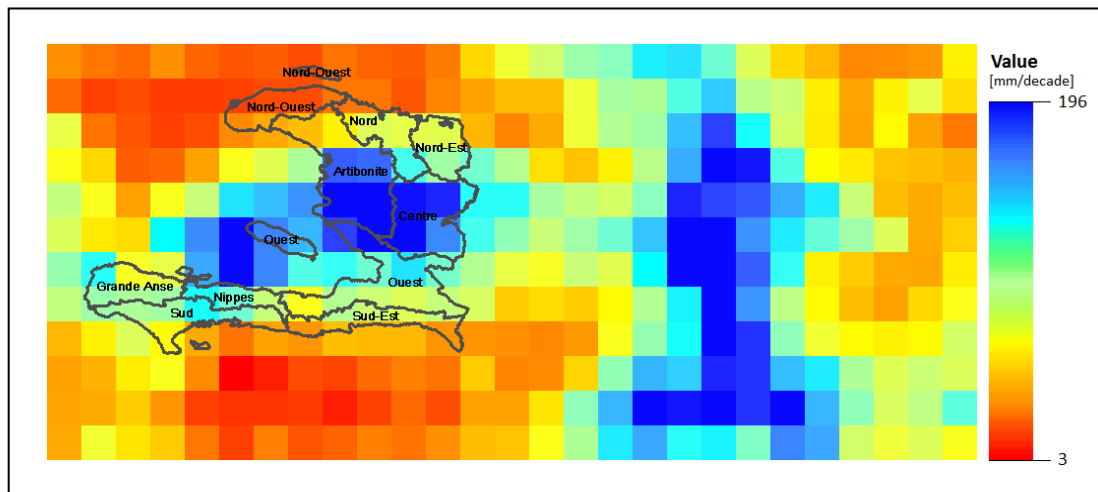
I dati giornalieri sono forniti dallo stesso *International Research Institute for Climate and Society* e si riferiscono al periodo dall'1 gennaio 2002 fino al 31 ottobre 2011 e hanno la stessa risoluzione di quelli decadali.

Per trasferire le precipitazioni dalle celle ai nodi (o ai dipartimenti), si assegna ad ogni bacino la precipitazione calcolata con l'algoritmo *nearest neighbor* eseguito sulle precipitazioni delle celle che coprono il bacino in considerazione. In questo modo si ottengono dei valori puntuali di precipitazione per le entità desiderate.

---

<sup>7</sup> Stime di precipitazione TRMM\_3B42, per approfondimenti consultare il sito <http://trmm.gsfc.nasa.gov/>

<sup>8</sup> Disponibile all'indirizzo <http://earlywarning.usgs.gov/fews/>

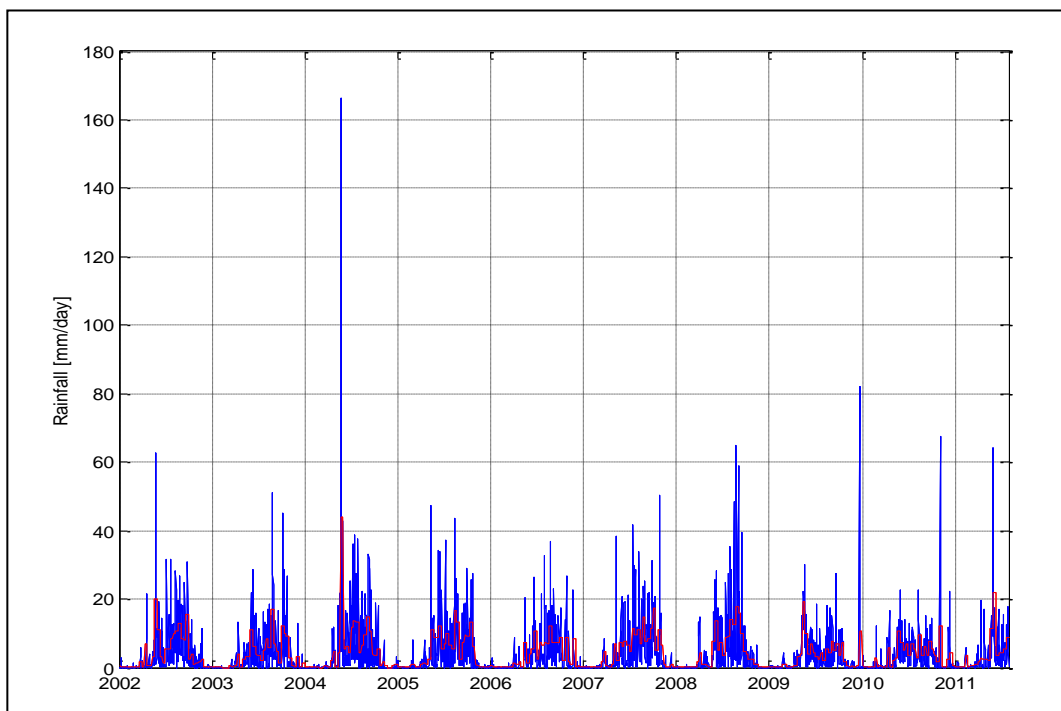


**Figura 3.3:** Rappresentazione dei dati di precipitazione decadali forniti dall'*International Research Institute for Climate and Society*. Si riporta soltanto Haiti. I dati sono relativi alla terza decade di luglio 2011. Ogni immagine è composta da 12x27 celle, ognuna delle quali di lato di 0.25 gradi (circa 28 km).

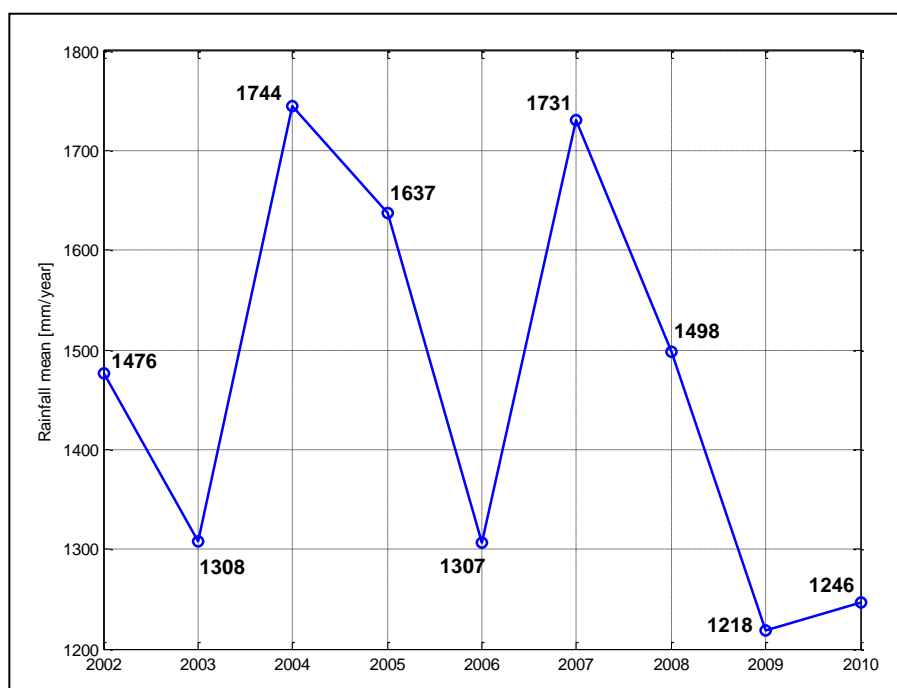
### 3.4.1 Statistiche di precipitazione

Le statistiche di precipitazione sono state calcolate sulla base sia dei dati giornalieri che di quelli decadali dell'*International Research Institute for Climate and Society*. Dalla **Figura 3.4**, che riporta entrambi i tipi di precipitazione, si può osservare che i dati decadali non rappresentano gli eventi estremi, in quanto riportano un valore mediato su dieci giorni. Inoltre è presente una significativa variabilità inter-annuale e intra-annuale, con un alternarsi di periodi particolarmente piovosi ad altri particolarmente secchi.

La media annua delle precipitazioni è di 1463 mm. L'anno più piovoso risulta essere il 2004 (**Figura 3.5**), con un totale di 1744 mm, mentre l'anno meno piovoso è il 2009, con 1218 mm, seguito dal 2010 con 1246 mm. Nel 2011, utilizzando i dati giornalieri, più aggiornati, al 31 ottobre si registra una precipitazione media su tutta Haiti di 1239 mm, mentre nel 2010 per lo stesso periodo si erano avuti solo 1053 mm.



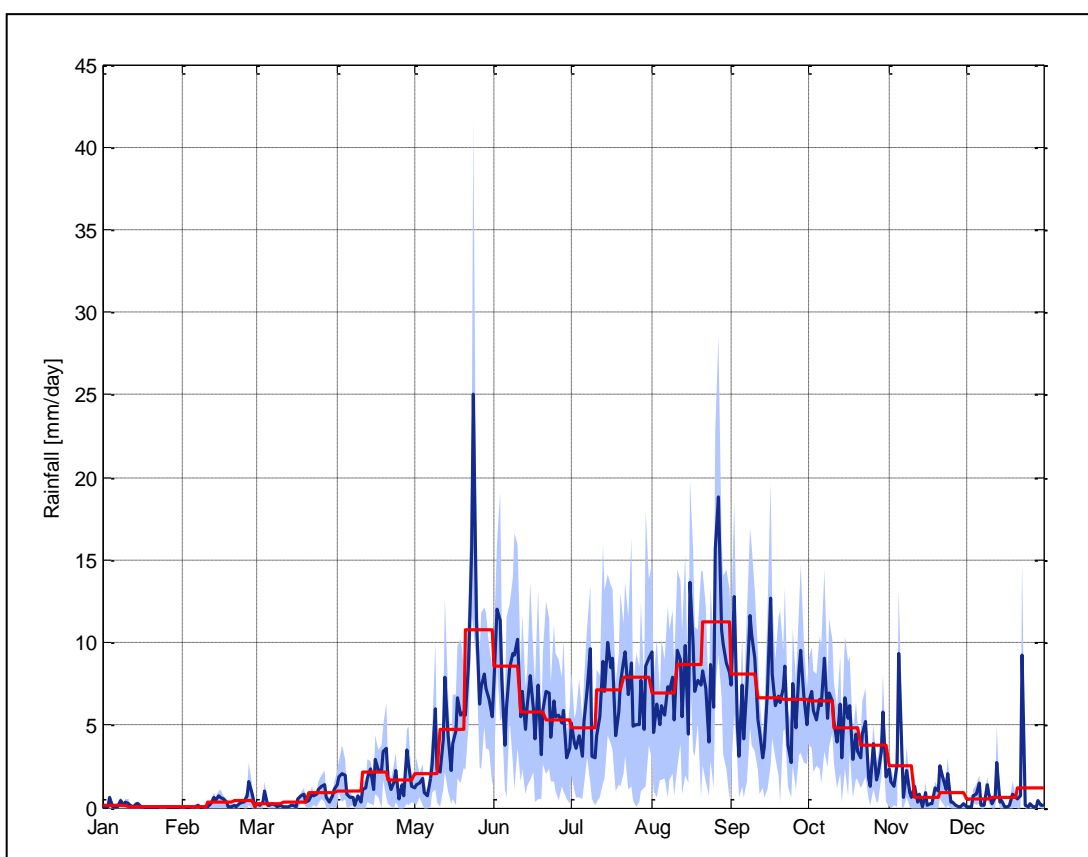
**Figura 3.4:** Andamento delle precipitazioni giornaliere (in blu) e di quelle decadali (in rosso) di Haiti per il periodo dal 1 gennaio 2002 al 31 luglio 2011 in mm/giorno.



**Figura 3.5:** Media annuale della precipitazione di Haiti per gli anni dal 2002 al 2010.

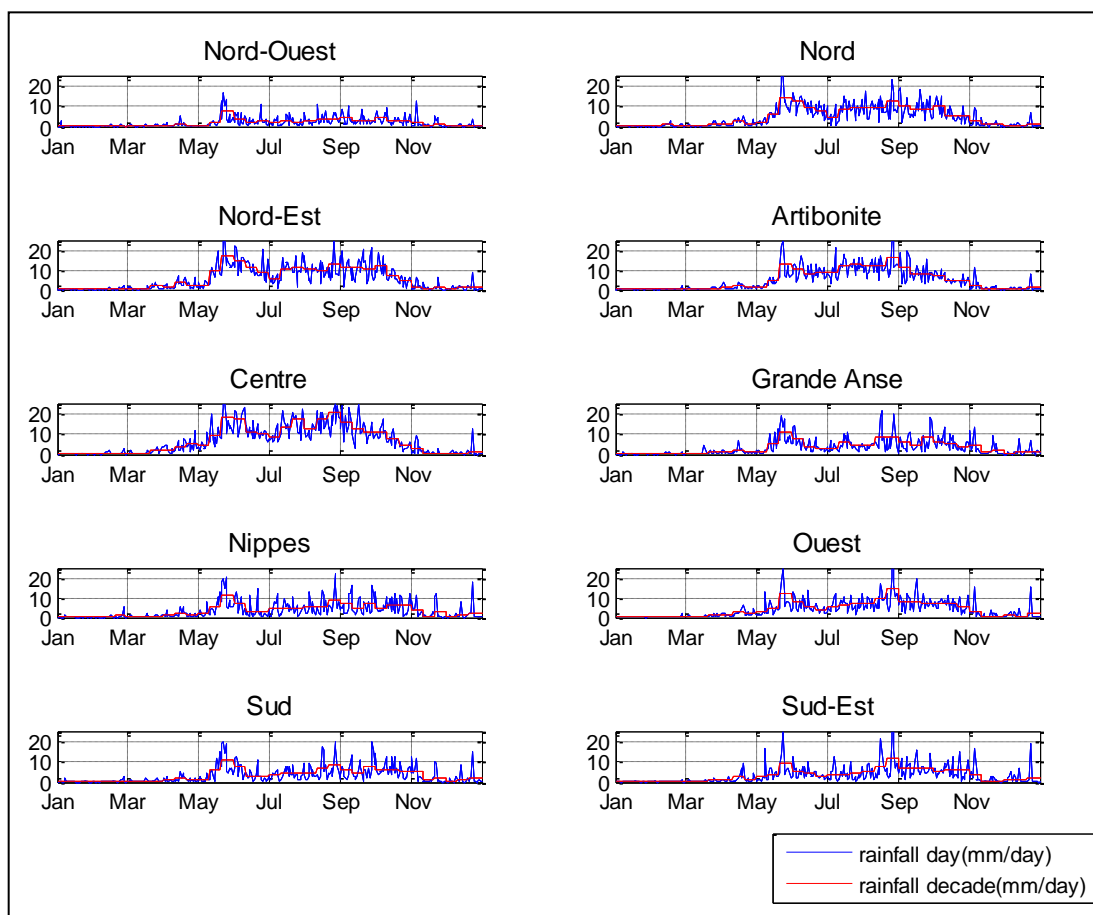


Il clima di Haiti è caratterizzato da due distinte stagioni piovose e da una stagione secca, nel periodo da dicembre a marzo (**Figura 3.6**). Durante la prima stagione piovosa, tra maggio e luglio, le precipitazioni possono superare anche i 160 mm/giorno (a maggio 2004), mentre durante la seconda, tra agosto e ottobre, generalmente si registra un valore più basso. Il periodo tra giugno e novembre corrisponde alla stagione degli uragani, che si abbattono sull'isola con maggiore probabilità tra la fine di agosto e settembre.



**Figura 3.6:** Andamento medio annuale della precipitazione di Haiti. In blu è riportato l'andamento della precipitazione media giornaliera, in rosso quello della precipitazione media decadale, mentre in azzurro è rappresentato l'intervallo di precipitazione giornaliera tra il 25-esimo percentile e il 75-esimo percentile.

Le precipitazioni più intense si registrano nella parte nord-orientale di Haiti (distretti di Centre e di Nord-Est), mentre il dipartimento meno piovoso è quello di Nord-Ouest (**Figura 3.7**). Per quest'ultimo, è presente una stagione particolarmente secca che va da dicembre a maggio. Per gli altri distretti è possibile osservare l'andamento stagionale delle precipitazioni e i due periodi di maggiore intensità di queste ultime. Dalla **Tabella 3.1** è possibile notare che alcuni dipartimenti hanno una precipitazione media giornaliera tra loro comparabile, come per esempio nella zona meridionale di Haiti, per i distretti di Sud, Sud-Est, Nippes e Grande Anse.



**Figura 3.7:** Andamento della precipitazione media dei dieci distretti di Haiti. In blu è riportato l'andamento della precipitazione media giornaliera, mentre in rosso è rappresentato l'andamento della precipitazione media decadale.

Distretto	Precipitazione media annua (mm/anno)	Precipitazione media giornaliera (mm/giorno)
<b>Nord-Ovest</b>	675	1.84
<b>Nord</b>	1714	4.69
<b>Nord-Est</b>	2018	5.52
<b>Artibonite</b>	1783	4.88
<b>Centre</b>	2513	6.88
<b>Grande Anse</b>	1212	3.31
<b>Nippes</b>	1235	3.38
<b>Ovest</b>	1506	4.12
<b>Sud</b>	1218	3.33
<b>Sud-Est</b>	1217	3.33

**Tabella 3.1:** Per ogni distretto sono riportate la precipitazione media annua (mm/anno) e quella media giornaliera (mm/giorno).

### 3.5 Correlazione spaziale della precipitazione

È ragionevole chiedersi se è possibile individuare delle aree con un regime pluviometrico apparentemente simile. Per verificare l'esistenza di un legame tra le precipitazioni dei diversi dipartimenti, si calcola la matrice di correlazione delle precipitazioni distrettuali giornaliere. Pertanto, per ogni distretto, si ottiene la correlazione delle sue precipitazioni rispetto a quelle degli altri dipartimenti.

I risultati vengono mostrati in **Figura 3.8**. Il dipartimento preso in esame è rappresentato con un colore verde scuro, mentre gli altri distretti sono colorati con una gradazione dal verde, che indica un'alta correlazione, al rosso, che indica una scorrelazione tra le precipitazioni delle aree.

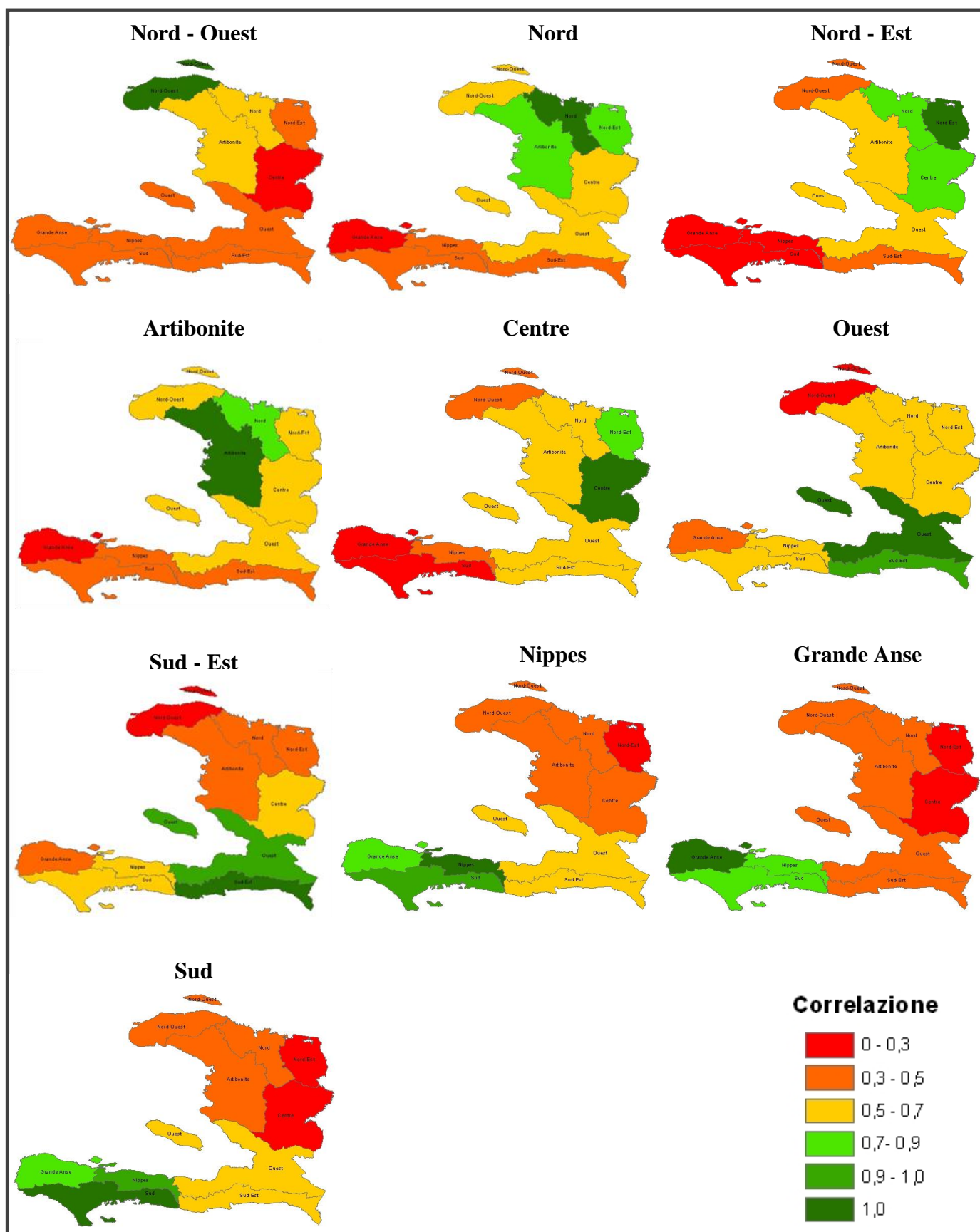


Figura 3.8: Mappe di correlazione tra la precipitazione di un dipartimento (in verde scuro) e quella dei distretti circostanti.

Il regime pluviometrico di Nord-Ouest non è simile a quello di nessuno degli altri distretti. Al nord, ad eccezione di Nord-Ouest, è possibile osservare una correlazione tra i distretti, ma non si riescono ad identificare dei gruppi in modo univoco. Per quanto riguarda la parte meridionale di Haiti, è possibile identificare delle zone con un'alta correlazione. In particolare si possono individuare due aree: la prima composta dai dipartimenti di Grande Anse, Nippes e Sud, e la seconda da quelli di Sud-Est e Ouest.

Con questa analisi viene confermata l'esistenza di un legame tra i regimi pluviometrici di alcuni dipartimenti. In particolare, è possibile notare che questa relazione è presente solo per aree contigue, ma il fatto che due distretti siano confinanti non implica che le loro precipitazioni siano correlate.

## VALIDAZIONE E PREVISIONE

In questo capitolo vengono effettuate le simulazioni dell'evoluzione dell'epidemia di colera ad Haiti integrate con le previsioni di precipitazione. Infatti quest'ultime costituiscono uno degli ingressi del modello SIRB (**Capitolo 2, Modello 2.4**). In tutte le simulazioni proposte con il modello epidemiologico, per il periodo della calibrazione (dal 23 ottobre 2010 fino al 28 maggio 2011) si utilizzano le precipitazioni decadali della *NASA-JAXA's Tropical Rainfall Measuring Mission*.

Innanzitutto si mostra la validazione del modello, dove le simulazioni sono eseguite inserendo come ingresso la precipitazione registrata dalla *NASA-JAXA's Tropical Rainfall Measuring Mission*. Successivamente, si presentano diversi metodi di previsione della precipitazione:

- Precipitazione media
- Precipitazione dell'anno tipico, ovvero l'anno osservato i cui valori meno si discostano dall'anno medio
- Metodo di previsione con modelli ARX
- Metodi di generazione puntuale della precipitazione di tipo poissoniano eseguita sia generando per distretti che per nodi

Con ognuno di questi metodi si eseguono delle previsioni su lungo periodo.

## 4.1 Validazione con precipitazioni reali

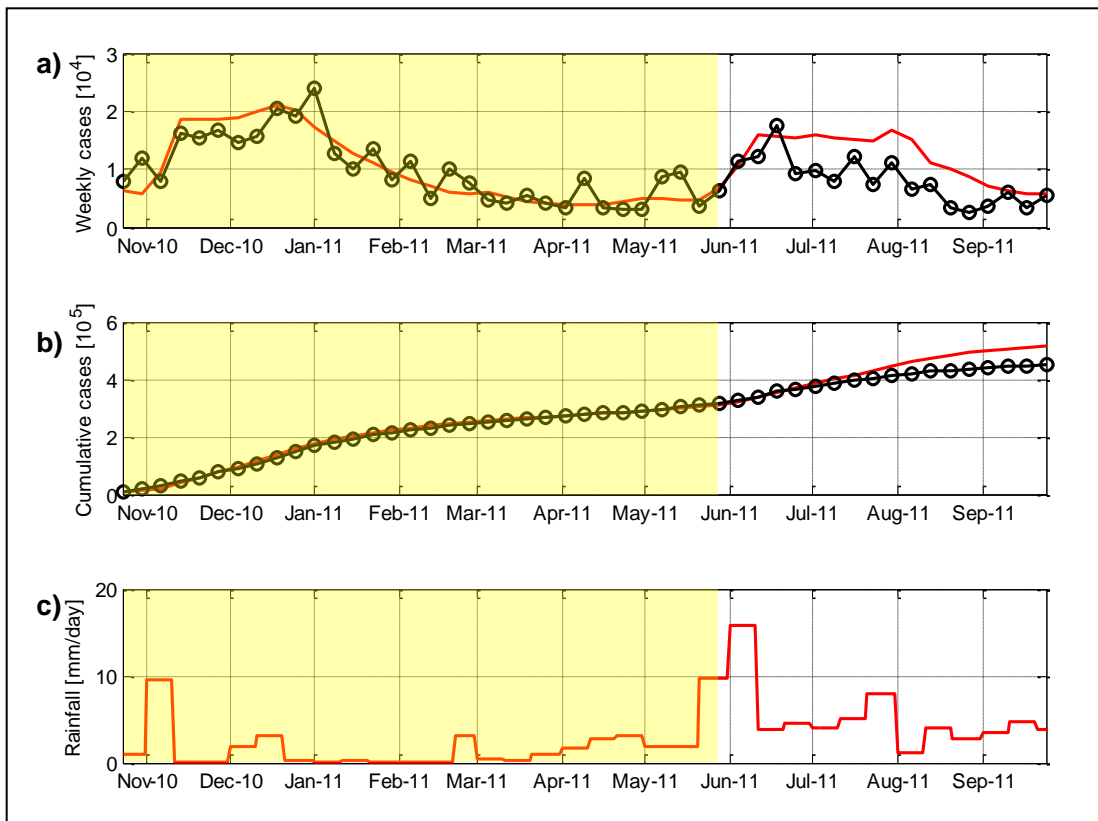
Avendo a disposizione i dati reali di precipitazione fino al 30 settembre 2011, con questi si realizza una prima simulazione del modello epidemiologico che può essere considerata come una validazione.

In **Figura 4.1** è possibile distinguere il periodo della calibrazione, evidenziato in giallo, da quello della validazione. Il primo dei tre grafici (**Figura 4.1a**) riporta l'andamento dei casi settimanali di colera sommati su tutto il territorio di Haiti, il secondo (**Figura 4.1b**) i casi cumulati e il terzo (**Figura 4.1c**) l'andamento delle precipitazioni.

Innanzitutto è possibile osservare che per il periodo relativo alla calibrazione (dal 23 ottobre 2010 al 28 maggio 2011), l'andamento dei casi simulati (linea rossa) rispecchia quello dei casi registrati (linea nera). In particolare, per quanto riguarda i casi cumulati (**Figura 4.1b**), al 28 maggio 2011 vengono registrati  $3.16 \cdot 10^5$  casi di colera, mentre ne vengono simulati  $3.12 \cdot 10^5$ . Considerando invece il periodo della validazione, che va dal 29 maggio 2011 al 30 settembre 2011, i casi di colera sono sovrastimati. Infatti, a fine settembre 2011 si registrano nella realtà  $4.53 \cdot 10^5$  casi cumulati di colera, contro i  $5.18 \cdot 10^5$  che ne vengono previsti. Con questo metodo si riesce a descrivere l'incremento dei casi settimanali di colera nel periodo tra maggio e giugno 2011, anche se il modello non riesce a riprodurre il rapido declino dei casi dopo il picco di metà giugno. Nonostante il fatto che il modello preveda 65 000 casi in più, che corrisponde ad un errore di stima di circa il 14%, in generale è preferibile sovrastimare i casi di colera piuttosto che il contrario, in modo da non sottovalutare la portata dell'evoluzione dell'epidemia.

Nella simulazione, escluso il periodo di forte diffusione iniziale, un aumento dei casi di colera è generalmente anticipato da un periodo di intensa precipitazione. Per esempio, per la prima decade di giugno si registra un picco di precipitazione e nelle due settimane seguenti si osserva un aumento dei casi registrati. Infatti, bisogna considerare che l'effetto dovuto alle precipitazioni sull'aumento del contagio non è istantaneo, in quanto il tempo di incubazione della malattia può durare qualche

giorno. Di conseguenza le simulazioni risultano essere attendibili e non distanti dalla realtà, ad indicare un'ottima prestazione del modello di previsione dell'epidemia di colera.



**Figura 4.1:** Calibrazione e validazione del modello epidemiologico (Capitolo 2, Modello 2.4).

a) Andamento dei casi settimanali di colera. In nero sono indicati i casi osservati e in rosso quelli simulati con la precipitazione reale per il periodo dal 23 ottobre 2010 al 30 settembre 2011.

b) Andamento dei casi di colera cumulati. In nero sono indicati i casi osservati e in rosso quelli simulati con la precipitazione reale.

c) Andamento delle precipitazioni reali mediate su tutta Haiti.

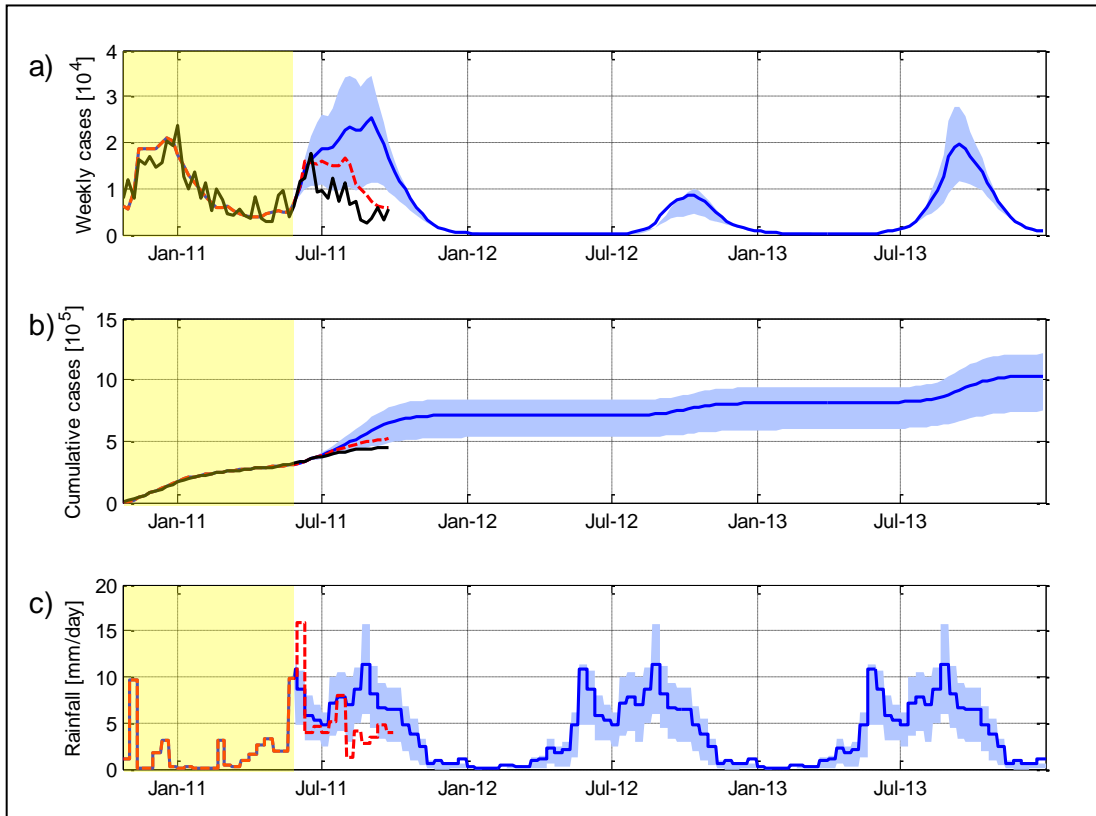
Lo sfondo giallo indica il periodo della calibrazione del modello, il bianco quello della validazione.



## 4.2 Simulazione con l'anno di precipitazione medio

Come primo metodo per proiettare in avanti le precipitazioni a partire dal 29 maggio 2011, si propone l'utilizzo delle precipitazioni medie, calcolate sulla base dei dieci anni di dati decadal (dal 2002 al 2011) forniti dall'*International Research Institute for Climate and Society*, opportunamente risolte alla scala dei nodi. Si esegue quindi una simulazione su lungo periodo a partire dal 23 ottobre 2010, utilizzando i dati reali di precipitazione fino al 28 maggio 2011 e quelli dell'anno medio dal 29 maggio 2011 al 31 dicembre 2013, in modo da poter prevedere eventuali risorgenze dell'epidemia. Per il periodo di previsione, oltre all'anno medio, si utilizzano anche il 25-esimo percentile delle precipitazioni di ogni nodo (quindi si impone che su tutta Haiti la precipitazione sia scarsa) e il 75-esimo percentile di ogni nodo (quindi si impone che piova molto e dappertutto).

Per quanto riguarda il periodo dal 29 maggio 2011 al 30 settembre 2011 è possibile confrontare direttamente simulazione e dati. Anche nella simulazione effettuata utilizzando la precipitazione dell'anno medio, si riesce a prevedere un aumento dei casi di colera nel periodo maggio-giugno (**Figura 4.2a**). Nei tre mesi successivi i casi simulati sono però fortemente sovrastimati rispetto a quelli reali. Questa differenza è causata dalla maggiore intensità delle precipitazioni medie rispetto a quelle reali. Infatti, sebbene l'andamento dell'anno medio preveda un aumento di precipitazione verso agosto-settembre, il 2011 è risultato essere un anno particolarmente secco, in cui non è presente la seconda stagione piovosa tipica di Haiti. In particolare, si osservi come nei due mesi presi in considerazione la precipitazione relativa al 2011 sia inferiore anche rispetto al primo quartile delle precipitazioni (**Figura 4.2c**). Questo fatto può rappresentare una possibile spiegazione della diminuzione dei casi di colera registrati rispetto a quelli simulati. Infatti nella simulazione, in corrispondenza del picco di precipitazione di fine agosto descritto dall'anno medio, è previsto un ulteriore aumento dei casi di colera settimanali.



**Figura 4.2:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) con l'anno di precipitazione medio.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, con una linea tratteggiata rossa quelli simulati utilizzando i dati di precipitazione reali e in blu quelli che utilizzano la precipitazione mediata sui dieci anni. Infine, la banda azzurra rappresenta l'intervallo dei casi simulati utilizzando il 25-esimo percentile e il 75-esimo percentile delle precipitazioni per ogni nodo.

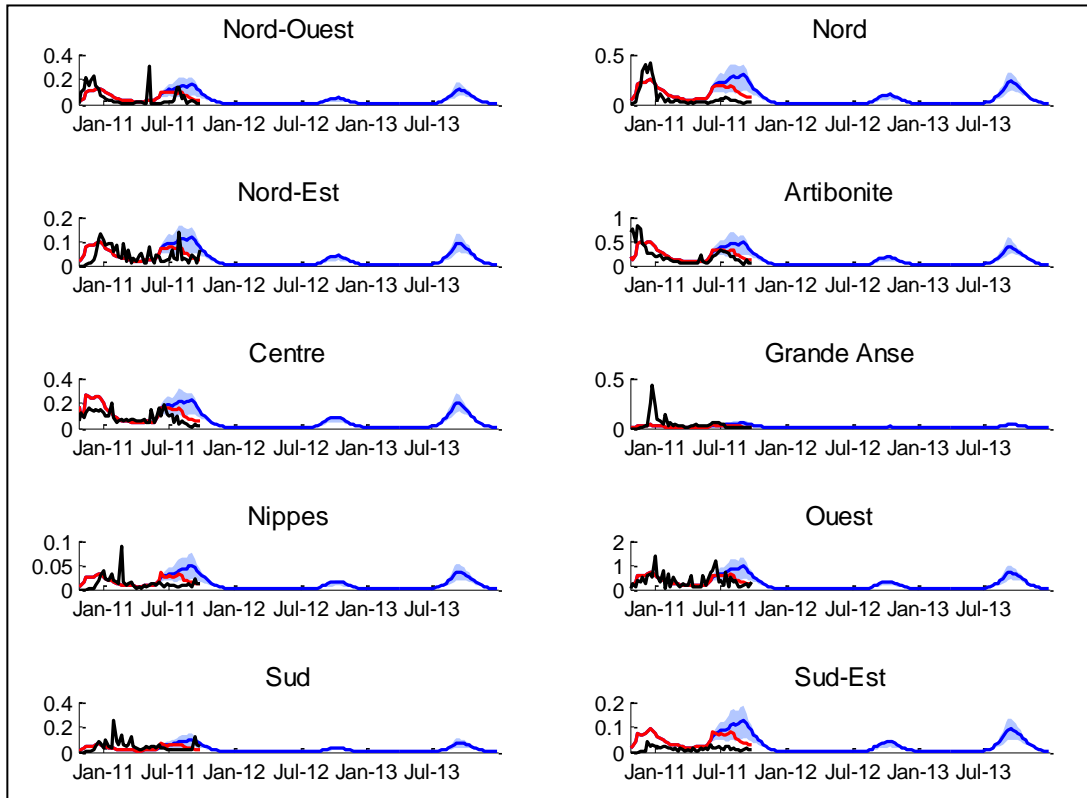
b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In rosso è rappresentato l'andamento della pioggia decadale secondo i dati della *NASA-JAXA's Tropical Rainfall Measuring Mission*, mentre in blu quello delle precipitazioni medie, usate per la simulazione. La banda azzurra rappresenta la media su Haiti del primo quartile e del terzo quartile delle precipitazioni per ogni nodo.

Lo sfondo giallo individua il periodo della calibrazione.

Confrontando i casi cumulati registrati e quelli simulati con la precipitazione media al 30 settembre 2011 (**Figura 4.2b**), si hanno rispettivamente  $4.53 \cdot 10^5$  e  $6.47 \cdot 10^5$  casi, quindi il modello sovrastima di circa 194 000 casi i casi di colera cumulati, con un errore di stima del 43%. Per quanto riguarda il 25-esimo percentile si prevedono  $4.85 \cdot 10^5$  casi cumulati (errore di stima del 7%) e per il 75-esimo percentile  $7.72 \cdot 10^5$  casi (errore di stima del 70%). Le simulazioni effettuate con il 25-esimo percentile restituiscono un errore sui casi cumulati minore, ma non simulano correttamente le variazioni dei casi settimanali di colera. Si comprende quindi l'importanza di avere a disposizione delle previsioni di precipitazione, poiché risultano essere strettamente collegate al risultato della simulazione.

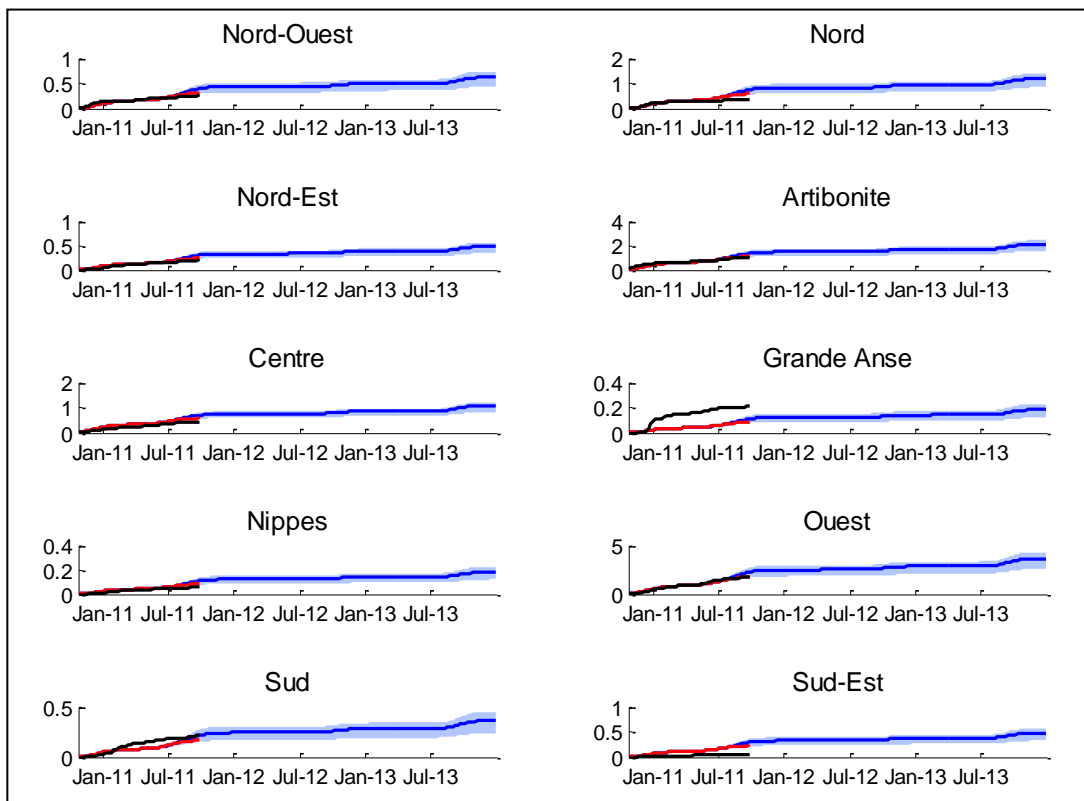
Venendo al periodo di previsione futura è possibile mettere in evidenza le seguenti caratteristiche. In corrispondenza con la stagione secca (da novembre 2011 a aprile 2012), l'epidemia sembra estinguersi, con un centinaio di casi a settimana in tutta Haiti. Si prevede una risorgenza dei casi di colera nel periodo tra agosto e novembre 2012, dopo le due stagioni delle piogge, con un picco di 8 726 casi settimanali previsto per il 13 ottobre 2012 (**Figura 4.2a**). Il contagio risulta essere contenuto rispetto al 2011, in cui si è registrato un picco di 23 864 casi settimanali reali l'1 gennaio 2011 e se ne erano addirittura previsti 25 232 per il 3 settembre 2011. Secondo la simulazione effettuata, per l'orizzonte di previsione, si verificherà anche una terza risorgenza della malattia nel periodo tra luglio e novembre 2013. Infatti, stabilendo che la perdita di immunità avviene mediamente dopo 3 anni, chi aveva contratto il colera nel 2010, da settembre 2013 ritorna mediamente a essere nuovamente suscettibile, ripopolando così il bacino delle persone potenzialmente infettabili. Questo fatto è ben visibile in **Figura 4.2a**, dove per i casi di colera settimanali simulati con l'anno di precipitazione medio è previsto un ulteriore picco di 19 605 casi per il 14 settembre 2013. Il modello, non considerando meccanismi di immunità latente, non distingue tra chi contrae la malattia per la prima volta e quelli che hanno già contratto la malattia in passato. Si noti come anche in questo caso la diffusione dell'epidemia sia immediatamente successiva al periodo delle piogge, ad indicare come queste influenzano il tasso di contagio.



**Figura 4.3:** Casi settimanali di colera per ogni distretto simulati col modello epidemiologico (Capitolo 2, Modello 2.4) che utilizza come ingresso l'anno di precipitazione medio. Sull'asse delle ordinate sono riportati i casi di colera divisi per  $10^4$ . Si osservi che le scale delle ordinate sono diverse per ogni grafico, in modo da consentire un confronto tra i casi reali e quelli simulati. In nero sono raffigurati i casi reali registrati, in rosso i casi simulati utilizzando le precipitazioni reali fino al 30 settembre 2011 e in blu i casi previsti con le precipitazioni relative all'anno medio. La fascia azzurra indica l'intervallo dei casi simulati con il primo quartile e il terzo quartile della precipitazione per ogni nodo.

In **Figura 4.3** si riportano le simulazioni dei casi settimanali registrati per ogni distretto. Per i dipartimenti di Ouest e per quello di Artibonite, che sono i distretti più popolati (in cui è presente rispettivamente circa il 36% e il 16% della popolazione totale), la previsione rispecchia le osservazioni relative a tutta Haiti (**Figura 4.2**). Infatti, la somma della popolazione presente in questi due distretti rappresenta gran parte della popolazione di Haiti ed è quindi corretto che sia presente una corrispondenza tra l'andamento dei casi di colera di questi dipartimenti e quelli sommati su tutta Haiti.

Osservando la **Figura 4.4**, è possibile notare come la simulazione effettuata con la precipitazione reale (in rosso) e con l'anno di precipitazione medio (in blu) rispecchi l'andamento dei casi cumulati reali, oltre che per i distretti di Ouest e Artibonite, anche per quelli di Nord-Ouest, Nord-Est, Nippes e Sud. Invece per i distretti Nord, Centre e Sud-Est, la simulazione sovrastima i casi cumulati di colera, mentre per Grande Anse li sottostima. Comunque il modello descrive in modo soddisfacente la diffusione dell'epidemia di colera per i sei distretti in cui è presente circa il 73% della popolazione totale. Utilizzare previsioni di precipitazione per analizzare in anticipo l'evoluzione dell'epidemia, pare introdurre incertezze quantitativamente accettabili, in quanto si riscontra in generale una sovrastima rispetto ai casi osservati.

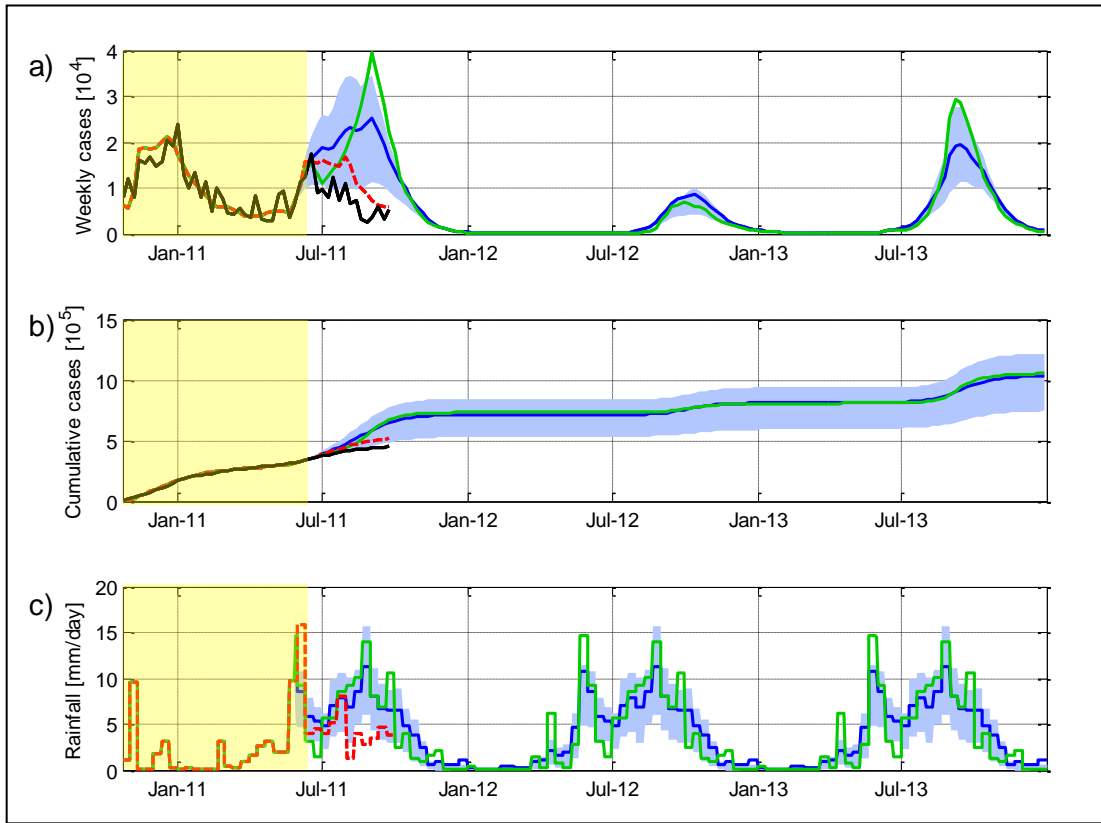


**Figura 4.4:** Casi cumulati di colera per ogni distretto simulati col modello epidemiologico (Capitolo 2, Modello 2.4) che utilizza come ingresso l'anno di precipitazione medio. Sull'asse delle ordinate sono riportati i casi di colera divisi per  $10^5$ . Si osservi che le scale delle ordinate sono diverse per ogni grafico, in modo da consentire un confronto tra i casi reali e quelli simulati. In nero sono raffigurati i casi reali registrati, in rosso i casi cumulati simulati utilizzando le precipitazioni reali fino al 30 settembre 2011 e in blu i casi previsti con le precipitazioni relative all'anno medio. La fascia azzurra indica l'intervallo dei casi cumulati simulati con il primo quartile e il terzo quartile della precipitazione per ogni nodo.

### 4.3 Simulazione con la precipitazione dell'anno tipico

In questo paragrafo, invece della precipitazione dell'anno medio, per il periodo dal 29 maggio 2011 al 31 dicembre 2013 si propone di utilizzare la precipitazione relativa all'anno tipico come ingresso nel modello di previsione dell'epidemia di colera. L'anno tipico è l'anno osservato i cui valori meno si discostano dall'anno medio. Studiando i dati decadalni forniti dall'*International Research Institute for Climate and Society* per gli anni dal 2002 al 2010, è possibile individuare il 2002 come anno tipico. Infatti, la differenza tra la precipitazione del 2002 e quella dell'anno medio presenta lo scarto minore. Inoltre anche le medie sono le più vicine, precisamente la media annua delle precipitazioni risulta essere di 1463 mm, mentre per il 2002 è di 1476 mm. Per la parte relativa alla calibrazione si utilizzano i dati della *NASA-JAXA's Tropical Rainfall Measuring Mission*, mentre per la previsione si usa l'anno tipico di precipitazione (e quindi i dati relativi al 2002). I risultati ottenuti con questo metodo vengono rappresentati in **Figura 4.5**.

L'andamento delle precipitazioni dell'anno tipico (**Figura 4.5c**) presenta dei picchi più alti e dei minimi più accentuati rispetto a quelli dell'anno medio, causando una risposta notevolmente diversa della malattia. Infatti si nota subito un picco di casi settimanali di colera a inizio settembre 2011 (**Figura 4.5a**). Entrando più nel dettaglio, per le prime due decadi del mese di giugno l'andamento delle precipitazioni dell'anno tipico segue meglio l'andamento delle precipitazioni registrate nella realtà rispetto all'anno medio, ma, per quanto riguarda la terza decade di giugno, la media della precipitazione giornaliera dell'anno tipico è di 1.37 mm/giorno, mentre per il 2011 risulta essere di 4.54 mm/giorno. Quindi il fatto che per il mese di giugno l'andamento dei casi di colera simulati con la precipitazione dell'anno tipico segua quello dei casi reali non deve ingannare, dato che si tratta di due ingressi di precipitazioni differenti. Nel periodo tra luglio e settembre 2011 l'andamento della precipitazione dell'anno tipico presenta un forte incremento e risulta essere superiore per ogni decade non solo rispetto alla precipitazione reale, ma anche rispetto a quella media, rimanendo però, ad eccezione della terza decade di settembre, sempre inferiore al 75-esimo percentile.



**Figura 4.5:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con la precipitazione dell'anno tipico.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera reali, con una linea tratteggiata rossa quelli simulati utilizzando i dati di precipitazione reali, in verde i casi ottenuti con la precipitazione dell'anno tipico e in blu quelli che utilizzano la precipitazione mediata sui dieci anni. Infine la banda azzurra rappresenta l'intervallo dei casi simulati utilizzando il 25-esimo percentile e il 75-esimo percentile della precipitazione per ogni nodo.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su Haiti. In rosso è rappresentato l'andamento della pioggia decadale secondo i dati registrati dalla *NASA-JAXA's Tropical Rainfall Measuring Mission*, in verde è riportato l'andamento delle precipitazioni dell'anno tipico (che corrisponde al 2002) e in blu quello delle precipitazioni dell'anno medio, usate per la simulazione. La banda azzurra rappresenta la media del primo quartile e del terzo quartile delle precipitazioni.

Lo sfondo giallo indica il periodo della calibrazione.

Questo provoca una rapida diffusione dell'epidemia, con un picco di casi settimanali pari a 39 449 previsto per il 3 settembre 2011. Si noti come questo massimo si verifichi poco dopo il picco di precipitazione della terza decade di agosto, ad indicare come il modello risponda sensibilmente alle variazioni di precipitazione. Un aumento così repentino dei casi di colera potrebbe essere spiegato o dal fatto che maggiori fluttuazioni nelle precipitazioni causano una risposta più impulsiva nella diffusione della malattia, o da una diversa distribuzione delle precipitazioni nei nodi. Infatti, un incremento dei casi di colera settimanali potrebbe essere dovuto anche al fatto che sia piovuto di più nei bacini più popolati. Analizzando il rapporto tra la precipitazione e la popolazione di uno stesso nodo per il periodo di fine agosto 2011, non si riesce però ad individuare una diversa distribuzione della precipitazione nei nodi rispetto al caso dell'anno medio. Ad esempio, la precipitazione nel nodo di Port-au-Prince, che raccoglie circa il 25% della popolazione di Haiti, risulta essere nel caso dell'anno tipico comunque inferiore a quella che si verifica considerando il 75-esimo percentile della precipitazione osservata. Se ne deduce che molto probabilmente il picco di casi di colera simulato utilizzando l'anno tipico di precipitazione sia dovuto alla fluttuazione della precipitazione.

#### **4.4 Simulazione con la precipitazione generata con un generatore puntuale di precipitazione poissoniano**

Un raffinamento ulteriore per la generazione degli afflussi meteorici è quello di usare un generatore puntuale (alla scala dei distretti o dei bacini) di precipitazione basato su un processo di tipo poissoniano (*point rainfall Poisson's generator*). Partendo da un'analisi statistica della serie storica di precipitazione, questo metodo genera delle precipitazioni con le stesse proprietà statistiche di quelle osservate. Generalmente per queste applicazioni si utilizzano serie temporali superiori ai 100 anni, ma in questo caso, si hanno a disposizione solo 10 anni di dati di

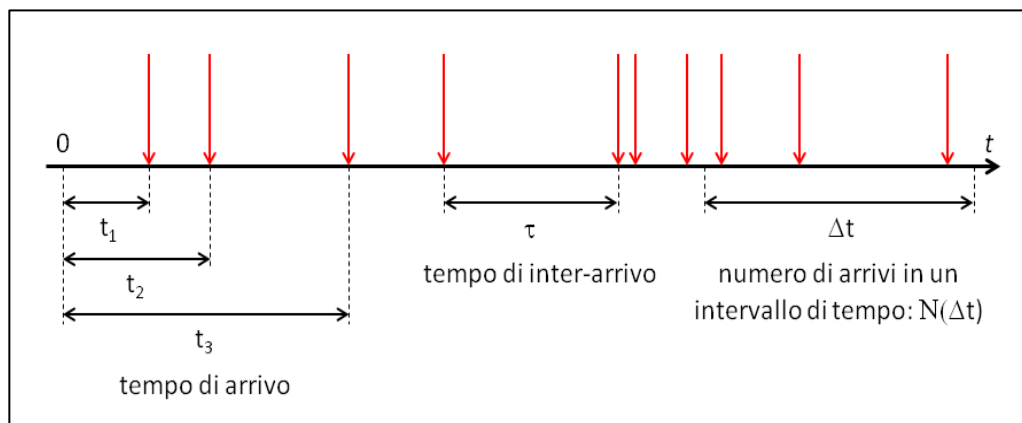


precipitazione. La generazione è di tipo puntuale e quindi non considera in modo esplicito le correlazioni spaziali.

In questo generatore si assume che l'intensità della precipitazione abbia una distribuzione esponenziale:

$$f(x) = \gamma e^{-\gamma x}$$

dove  $x$  è l'intensità della precipitazione, e  $1/\gamma$  il valore medio dell'intensità della precipitazione.



**Figura 4.6:** Schema delle grandezze caratteristiche in un processo di tipo poissoniano.

Le altre caratteristiche statistiche prese in considerazione sono (**Figura 4.6**):

1. Tempo di inter-arrivo, ovvero il tempo di attesa tra un evento e quello successivo ( $\tau$ )
2. Tempo di attesa di ogni arrivo a partire dall'istante iniziale ( $t_1, t_2, \dots, t_n$ )
3. Numero di arrivi in un prefissato intervallo di tempo ( $N(\Delta t)$ )

Da cui si possono formulare tre definizioni equivalenti:

1. Il tempo di inter-arrivo ha una distribuzione di tipo esponenziale, con valor medio  $1/\lambda$ :

$$f(\tau) = \lambda e^{-\lambda \tau}$$

2. Il tempo di attesa del  $k$ -esimo arrivo ha una distribuzione Gamma:

$$f(t_k) = \frac{(\lambda t_k)^{k-1}}{(k-1)!} \lambda e^{-\lambda t_k}$$

3. Il numero di arrivi in un determinato intervallo di tempo  $\Delta t$  segue la distribuzione di Poisson:

$$P[N(\Delta t) = n] = \frac{\lambda \Delta t^n}{n!} e^{-\lambda \Delta t}$$

Sebbene siano disponibili i dati di precipitazione anche relativi a parte del 2011, in questo tipo di generatore è possibile utilizzare soltanto gli anni completi, e quindi si lavora con soli nove anni di dati (2002-2010). È possibile utilizzare il generatore sia a partire dalla precipitazione di ogni distretto che da quella di ogni nodo.

Innanzitutto si devono calcolare le grandezze su cui si basa la generazione, che sono:

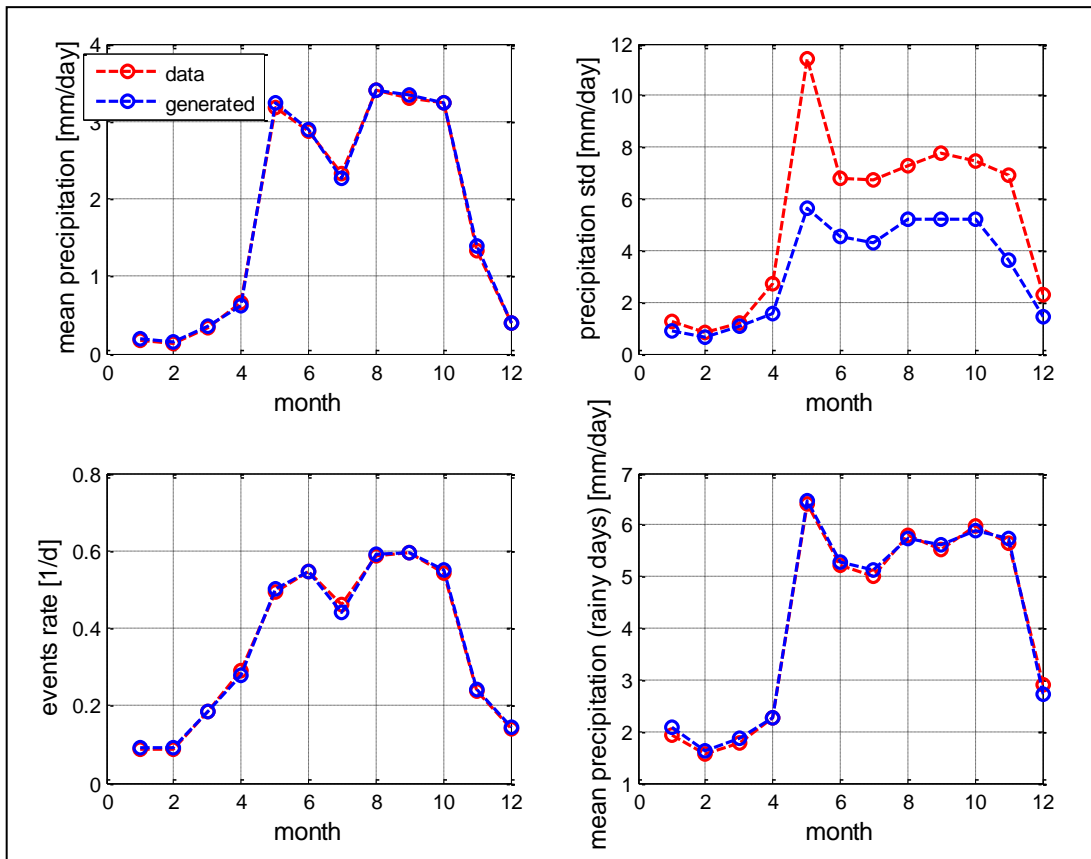
- Tasso di arrivo dei giorni piovosi in un mese ( $\lambda$ )
- Precipitazione media mensile (contando solo i giorni piovosi)

La procedura per stimare i parametri dei generatori è quella di far coincidere i valori teorici dei parametri dei generatori con quelli campionari. In particolare, per generare la precipitazione di un giorno per un determinato nodo (o distretto), la procedura è la seguente:

1. Si estrae un numero casuale da una distribuzione uniforme, per stabilire se in quel giorno piove oppure no. In particolare, piove se il numero casuale estratto dall'intervallo  $[0, 1]$  è minore o uguale a  $\lambda$  ( $\lambda = \frac{\text{numero di giorni in cui piove}}{\text{numero totale di giorni}}$ ). Se non piove, per quel giorno la precipitazione viene posta uguale a zero.
2. Se piove, si genera il valore della precipitazione. Anche in questo caso si estrae un numero casuale nell'intervallo  $[0, 1]$  da una distribuzione uniforme, che rappresenta la probabilità dell'evento. Il valore dell'intensità di precipitazione si calcola ipotizzando che quest'ultima abbia una distribuzione esponenziale.

Utilizzando i generatori si originano cento anni di precipitazione per ogni nodo o dipartimento. In **Figura 4.7** si riporta un esempio relativo alla generazione della

precipitazione del distretto Nord-Ouest. È presente un'ottima corrispondenza tra le precipitazioni reali (in rosso) e quelle generate (in blu) per quanto riguarda la precipitazione media, il tasso di accadimento e la precipitazione media dei giorni piovosi. Al contrario, il modello di generazione utilizzato non riesce a cogliere interamente la deviazione standard. Per rendere più aderente ai dati la precipitazione generata anche nella corrispondenza della deviazione standard, si potrebbero utilizzare modelli più complessi, ad esempio che tengano conto della variabilità interannuale delle precipitazioni [Porporato *et al.*, 2006].



**Figura 4.7:** Statistiche calcolate sui dati generati con il generatore puntuale di precipitazione di Poisson per il distretto di Nord-Ouest. In rosso sono rappresentate le statistiche reali e in blu quelle dei dati generati con il generatore di precipitazione.

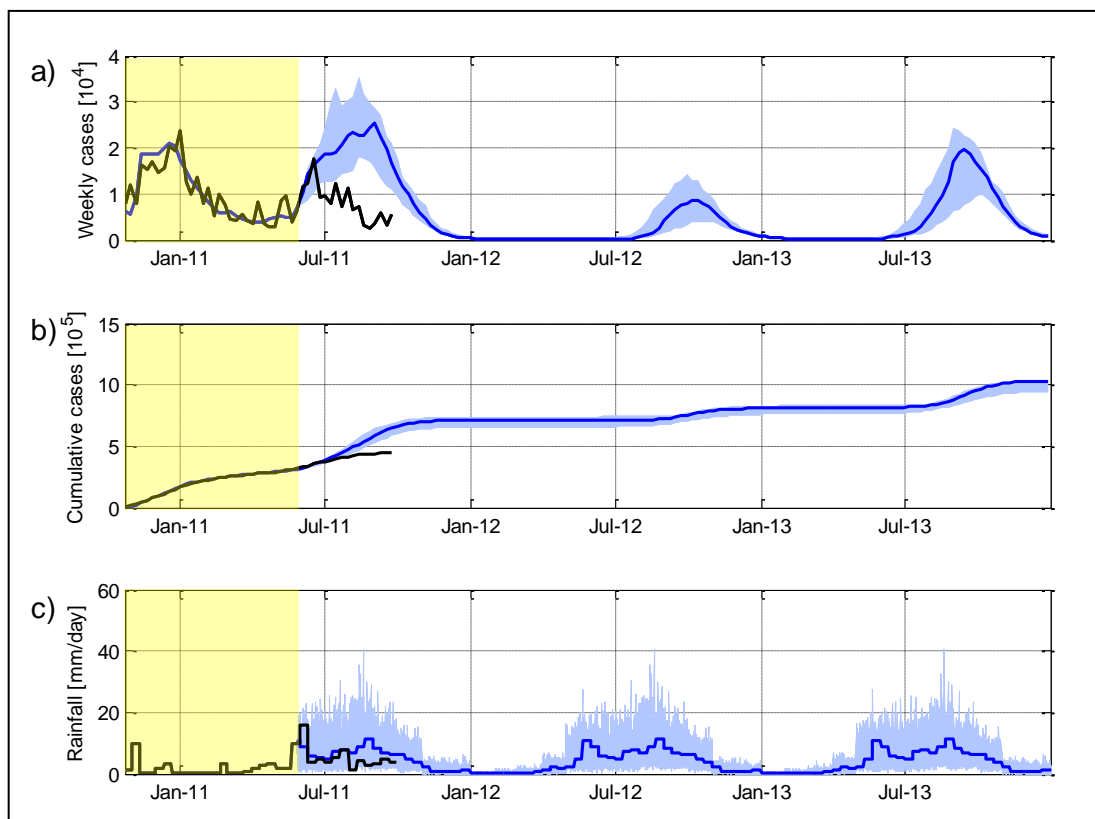
In conformità a queste statistiche si generano per ogni distretto (o bacino) cento anni di precipitazione da inserire come ingresso nel modello di colera. Dato che si vogliono effettuare simulazioni a lungo termine (fino al 31 dicembre 2013), è necessaria una sequenza di tre anni di dati di precipitazione simulati. Per creare questa sequenza, ogni anno originato con il generatore è seguito da un anno estratto casualmente tra i cento generati e così anche per il terzo (dove l'estrazione avviene sempre tra i cento anni generati). In questo modo si creano cento sequenze casuali di tre anni di precipitazione dagli anni generati dal generatore puntuale di tipo poissoniano.

#### **4.4.1 Generazione per dipartimenti disaggregati**

Inizialmente si originano i 100 anni di precipitazione basandosi sui nove anni di precipitazioni giornaliere dei distretti (2002-2010). Dopodiché, poiché il modello epidemiologico lavora sui nodi, viene effettuato il *downscaling* della pioggia dai distretti ai bacini, considerando la frazione di ognuno di questi ultimi presente in ogni distretto; in questo modo si modifica il dato per poterlo introdurre come ingresso nel modello di colera. Si esegue la simulazione con il modello SIRB (**Capitolo 2, Modello 2.4**) utilizzando la precipitazione reale dal 23 ottobre 2010 al 28 maggio 2011 e quella generata dal 29 maggio 2011 al 31 dicembre 2013 (**Figura 4.8 e Figura 4.9**).

Osservando il grafico relativo alle precipitazioni (**Figura 4.8c**), si nota immediatamente che le precipitazioni generate a partire da dati giornalieri (nel grafico sono riportate le medie sui nodi di ognuno dei cento anni generati) presentano una maggior fluttuazione giornaliera rispetto alle precipitazioni medie, calcolate su base decennale. Comunque, le previsioni dei casi di colera eseguite con la precipitazione originata dal generatore puntuale di precipitazione di Poisson seguono quella effettuata con la precipitazione media per tutto il periodo di simulazione (**Figura 4.8a e Figura 4.8b**). Questo fatto può essere spiegato ricordando che, come mostrato in **Figura 4.7**, la media mensile degli anni generati e quella calcolata sui

dati reali a disposizione corrispondono, così come il numero di eventi e la media delle precipitazioni nei giorni piovosi.



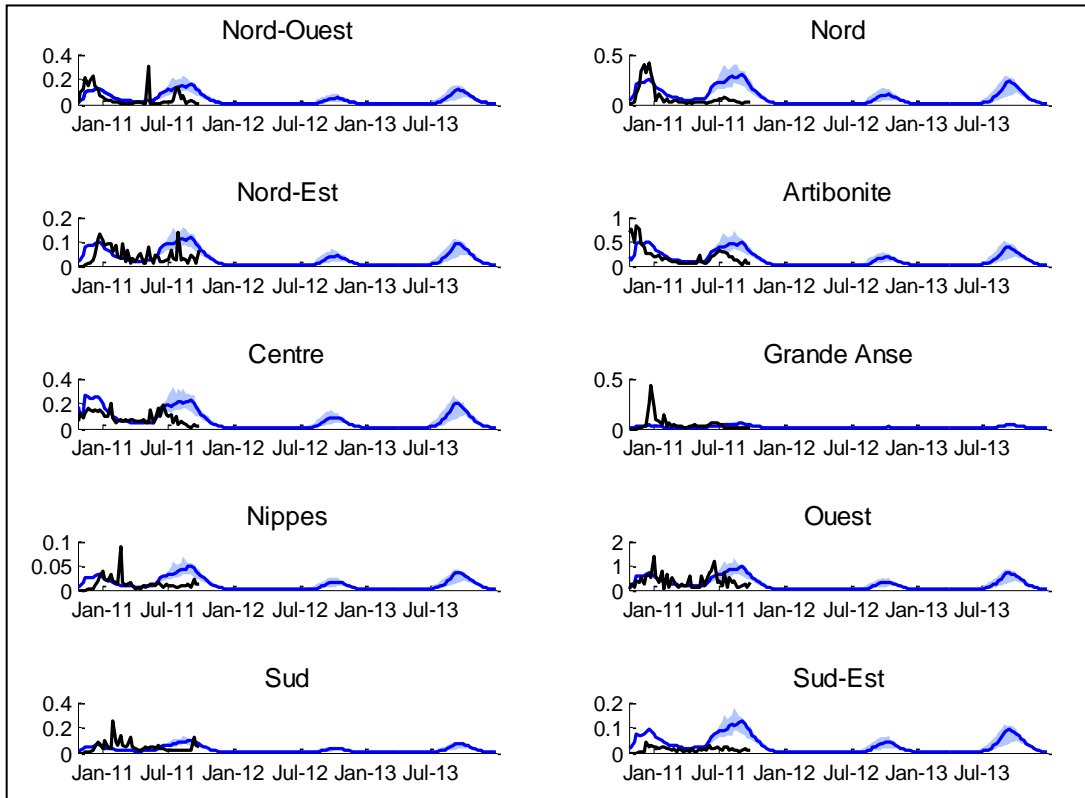
**Figura 4.8:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) effettuata con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione dei dipartimenti disaggregati.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, mentre in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale poissoniano a partire dalla precipitazione dei dipartimenti disaggregati.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In nero sono raffigurati i dati reali, mentre in blu sono indicate le precipitazioni medie. La fascia azzurra rappresenta l'intervallo tra le precipitazioni giornaliere minime e massime (mediate sui nodi per ognuno dei cento anni generati) generate con il generatore puntuale di precipitazione poissoniano a partire dalla precipitazione dei dipartimenti disaggregati.

Lo sfondo giallo indica il periodo della calibrazione.



**Figura 4.9:** Casi di colera settimanali per ogni distretto simulati con il modello epidemiologico (Capitolo 2, Modello 2.4) che utilizza come ingresso la precipitazione generata con il generatore puntuale di precipitazione di Poisson partendo dalla precipitazione giornaliera dei distretti disaggregati. Sull'asse delle ordinate i casi sono riportati divisi per  $10^4$ . In nero sono riportati i casi di colera registrati, in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dalla precipitazione dei dipartimenti disaggregati.

A fine settembre 2011 i casi cumulati di colera registrati sono 452710, quelli simulati con l'utilizzo dell'anno medio 647 220, mentre le previsioni effettuate basandosi sulla precipitazione generata con il generatore puntuale poissoniano prospettano 582 980 casi cumulati per lo scenario più ottimista e di 692 310 per lo scenario più catastrofico (**Figura 4.8b**). Il fatto che il modello sovrastimi i casi di colera è visto positivamente in quanto è meglio che si attivi una condizione di allerta che arrivare impreparati a un improvviso dilagare dell'epidemia.

Generando 100 anni di precipitazione si ottiene una previsione della varietà degli scenari possibili, che potrebbe essere ampliata generando un numero maggiore di anni (per esempio 1000 anni). Si ricorda infatti che il valore intrinseco di questa analisi risiede principalmente nella definizione di un gran numero di possibili scenari di precipitazione e di conseguenza anche epidemiologici, che in questo caso sono generati in base alle proprietà statistiche proprie del caso in questione. L'aderenza delle simulazioni ai dati reali rappresenta un risultato secondario rispetto al possibile utilizzo di questo strumento e, probabilmente, di difficile conseguimento con una base di dati così limitata. Grazie alle previsioni dei casi di colera simulati, si possono effettuare delle valutazioni di interventi a priori, da correggere con delle politiche di controllo in linea dell'epidemia.

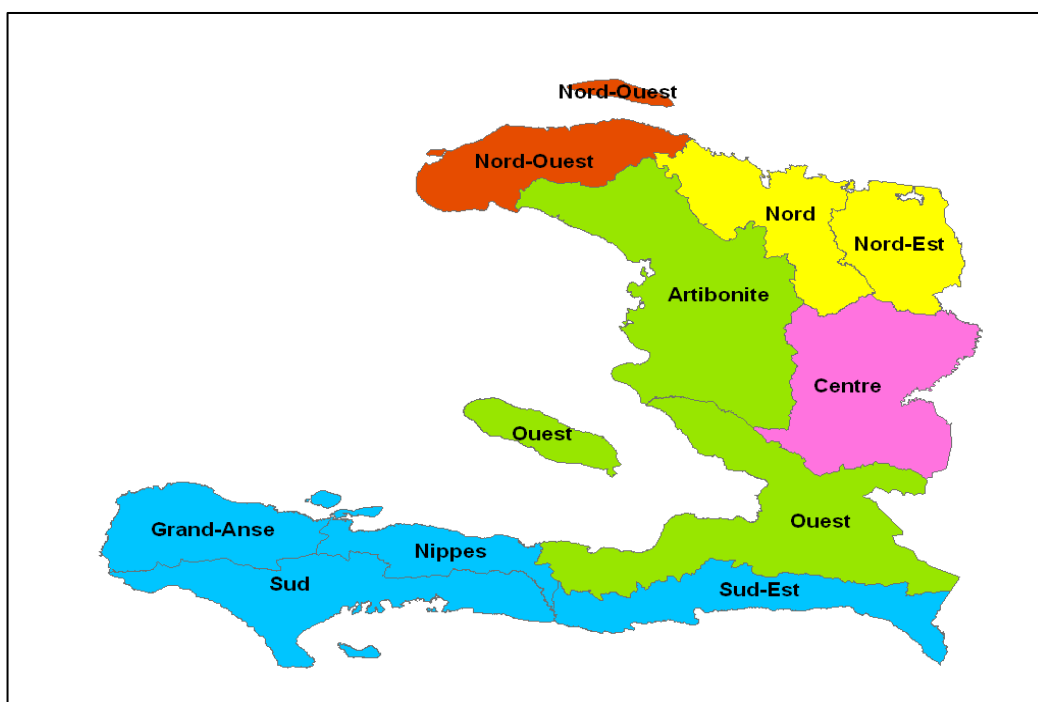
#### **4.4.2 Generazione per dipartimenti aggregati**

In precedenza i distretti sono stati considerati indipendenti tra loro. Tuttavia diverse aree di Haiti hanno un regime pluviometrico simile (**Figura 3.7**). Pertanto, si introducono delle aggregazioni tra distretti con caratteristiche simili nei *pattern* di precipitazione, definendo così un generatore di precipitazione per ogni gruppo di dipartimenti e non più per ogni singolo distretto. In riferimento agli andamenti medi mensili della precipitazione e alle osservazioni relative alla correlazione spaziale della precipitazione (**Capitolo 3.5**), il territorio di Haiti è stato suddiviso in cinque zone (**Figura 4.10**):

- Nord-Ouest
- Nord e Nord-Est
- Artibonite e Ouest
- Centre
- Sud-Est, Sud, Nippes e Grande Anse.

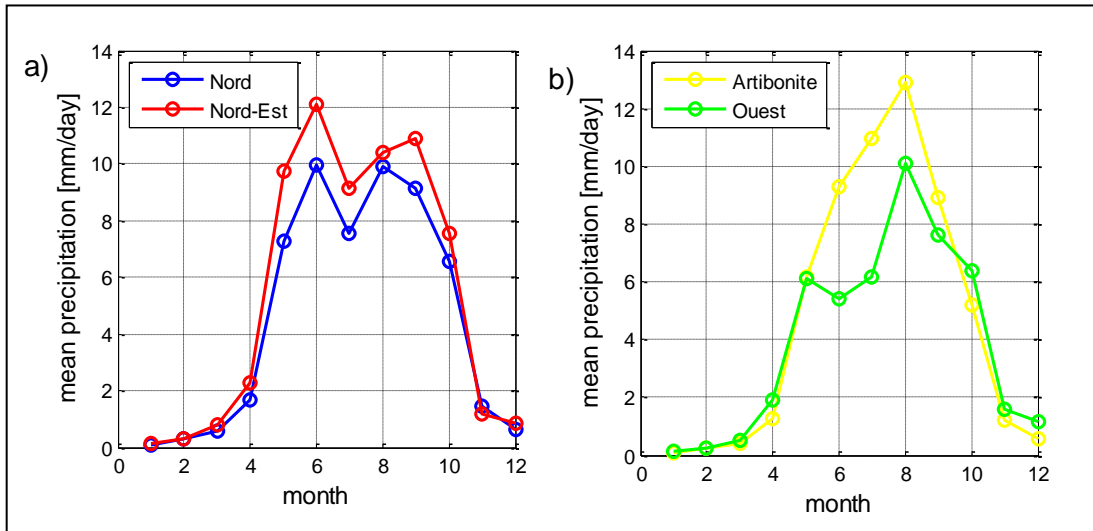
Il regime precipitativo del distretto di Nord-Ouest non è paragonabile a nessuno degli altri distretti, infatti qui si registrano le precipitazioni meno intense di Haiti; così anche per Centre che, al contrario, è il distretto in cui si riportano le

precipitazioni più intense. Di conseguenza questi due dipartimenti sono considerati singolarmente. Il secondo e il terzo gruppo, formato rispettivamente dai distretti di Nord e Nord-Est e da Artibonite e Ouest, presentano lo stesso andamento mensile delle precipitazioni. Come mostrato in **Figura 4.11a**, per il primo dei due gruppi si osserva un'ottima coincidenza; mentre per quanto riguarda **Figura 4.11b**, si è scelto di raggruppare Artibonite e Ouest, in quanto presentano medie paragonabili e hanno lo stesso andamento delle precipitazioni medie mensili con un picco di precipitazioni in agosto. Il gruppo composto da Sud-Est, Sud, Nippes e Grande Anse è formato da distretti con regime precipitativo paragonabile (**Figura 4.12**), con l'evidenza di due stagioni piovose. Inoltre, in riferimento alla **Tabella 3.1**, si osservi come anche le precipitazioni medie annue siano confrontabili (rispettivamente di 1217 mm/anno, 1218 mm/anno, 1235 mm/anno e 1212 mm/anno).



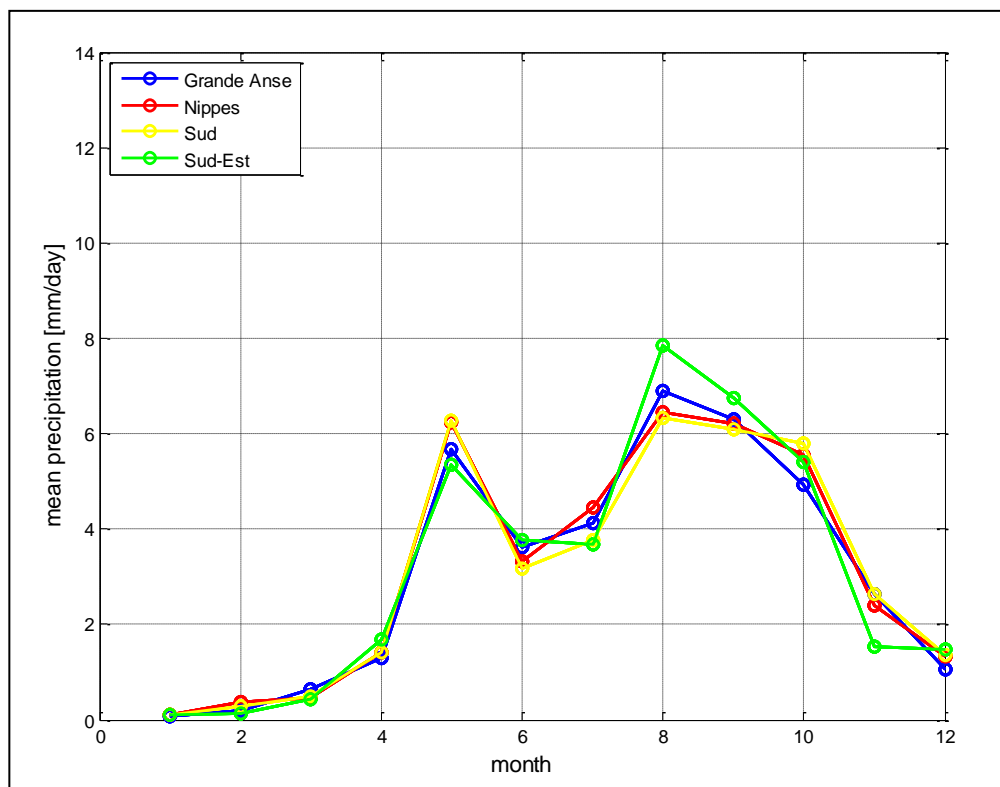
**Figura 4.10:** Suddivisione di Haiti in zone con precipitazioni comparabili.





**Figura 4.11:** Andamento delle precipitazioni medie mensili di:

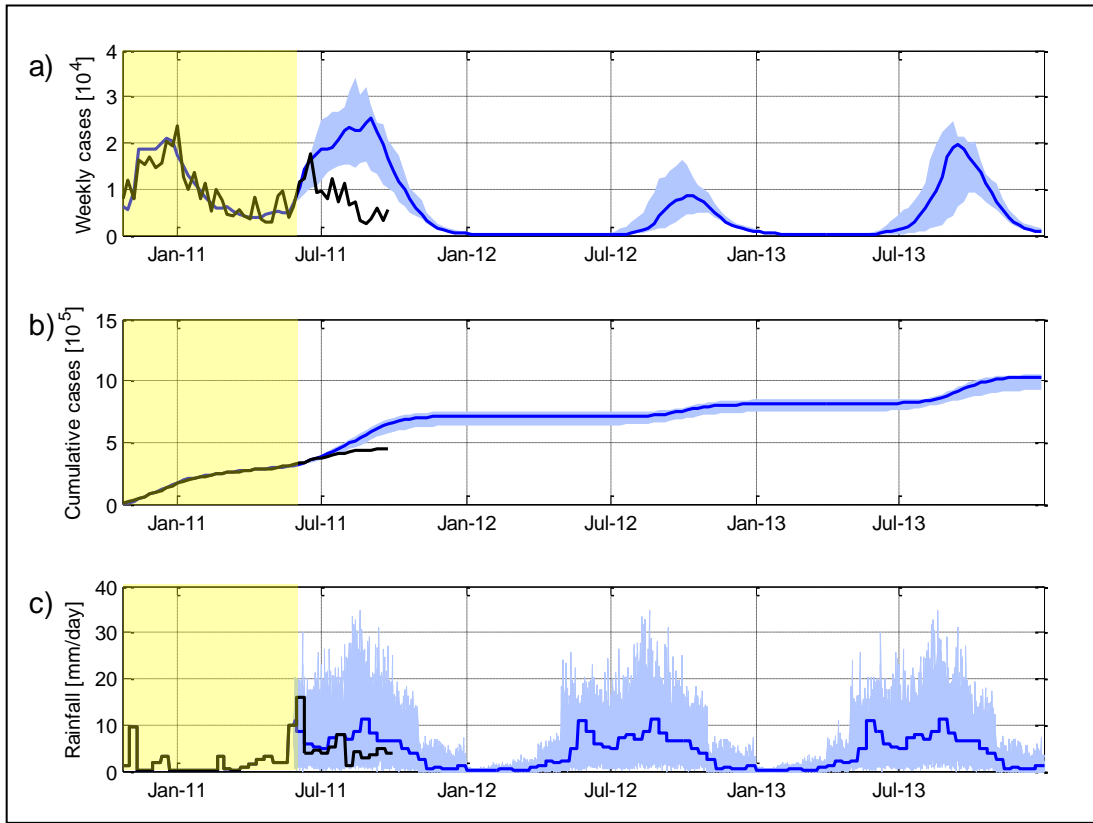
- a) Distretti di Nord (in blu) e Nord-Est (in rosso)
- b) Distretti di Artibonite (in giallo) e di Ouest (in verde)



**Figura 4.12:** Andamento delle precipitazioni medie mensili dei distretti di Grande Anse (in blu), Nippes (in rosso), Sud (in giallo) e Sud-Est (in verde).

Si calcola per ogni gruppo la media della precipitazione giornaliera per ogni giorno dall'1 gennaio 2002 al 31 dicembre 2010 sulla base dei dati giornalieri forniti dall'*International Research Institute for Climate and Society*. Si ottengono quindi cinque serie che vengono inserite come ingresso nel generatore puntuale di precipitazioni di Poisson, che calcola cento anni di precipitazione per ogni gruppo. Successivamente, ad ogni distretto, si assegnano le precipitazioni generate del gruppo corrispondente. Infine, poiché i cento anni di precipitazione generati si riferiscono ai distretti, si scala la precipitazione dai distretti ai nodi, sui quali lavora il modello. I risultati ottenuti con questo metodo sono rappresentati in **Figura 4.13**.

Rispetto al caso precedente, in cui la precipitazione era generata per ogni distretto, non si prevede il picco di casi settimanali a luglio 2011 (**Figura 4.13a**). Il motivo per cui non sono presenti sostanziali differenze neanche a scala di distretto (**Figura 4.9** e **Figura 4.14**) può essere imputato al fatto che il generatore puntuale di precipitazione di Poisson si basa su serie storiche, che quindi hanno già insite come informazione implicita la correlazione spaziale.



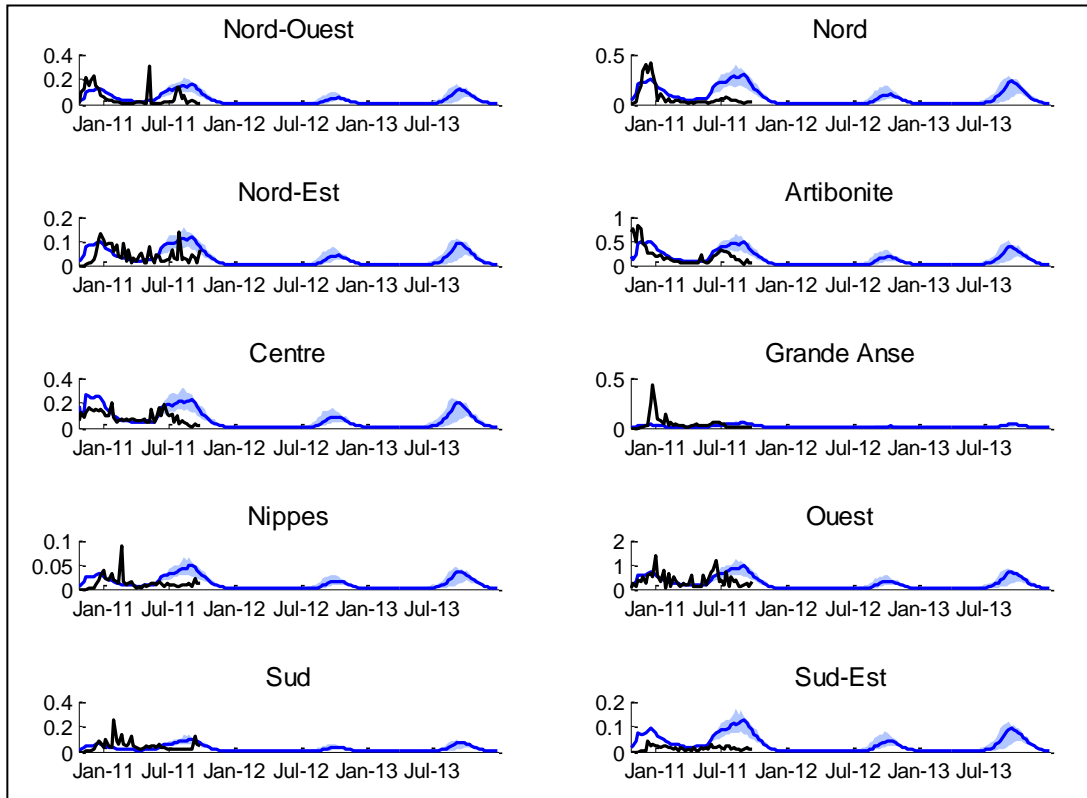
**Figura 4.13:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) effettuata con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei dipartimenti aggregati.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati e in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei dipartimenti aggregati.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In nero sono raffigurati i dati reali, mentre in blu sono indicate le precipitazioni medie. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi delle precipitazioni giornaliere (mediate sui nodi per ognuno dei cento anni generati) generate con il generatore puntuale di Poisson a partire dalla precipitazione dei dipartimenti aggregati.

Lo sfondo giallo indica il periodo della calibrazione.

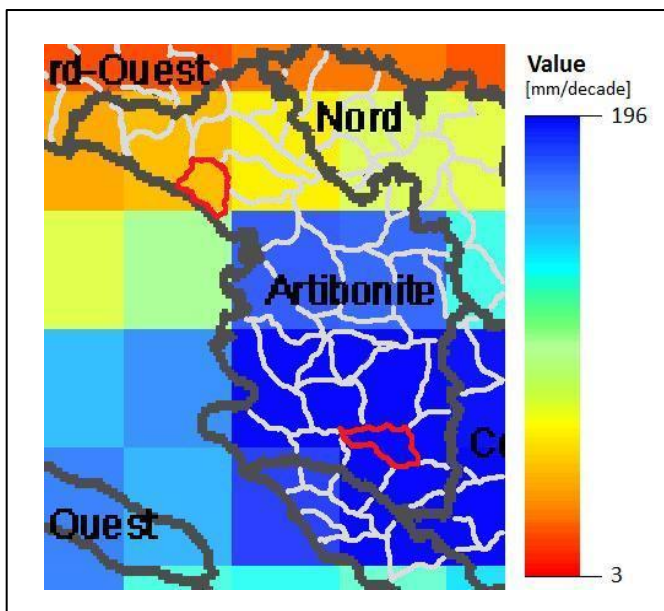


**Figura 4.14:** Casi di colera settimanali per ogni distretto simulati con il modello epidemiologico (Capitolo 2, Modello 2.4) che utilizza come ingresso la precipitazione generata con il generatore puntuale di precipitazione di Poisson partendo dalla precipitazione giornaliera dei distretti aggregati. Sull'asse delle ordinate i casi sono riportati divisi per  $10^4$ . In nero sono riportati i casi di colera registrati, in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei dipartimenti aggregati.

## 4.5 Simulazione con la precipitazione generata con un generatore puntuale di precipitazione poissoniano a partire dai bacini

In questo paragrafo si propone una generazione della precipitazione è che si basa sui nodi, in quanto sono le entità su cui lavora il modello di diffusione del colera. In questo caso, si preserva, nel *downscaling* dei dati di partenza, la precipitazione

specifica nel territorio afferente al nodo, cosa non garantita nel caso in cui si utilizzi un generatore a livello di dipartimenti. Infatti, quando si scala la precipitazione sui distretti, ad ognuno di essi si attribuisce un unico valore di precipitazione puntuale, su cui si basa il generatore poissoniano. Dopo aver generato cento anni di precipitazione per ogni dipartimento, si scala il dato di pioggia sui bacini. In particolare, ad ogni bacino completamente all'interno di un distretto viene attribuito il valore di precipitazione del distretto stesso, mentre per i bacini a cavallo di più distretti il valore della precipitazione è calcolato con una media pesata rispetto all'area all'interno di ogni dipartimento. Di conseguenza, poiché il valore della precipitazione è puntuale e non areale, quando si passa dai distretti ai bacini potrebbe essere attribuito lo stesso valore di precipitazione a dei bacini che invece registrano precipitazioni notevolmente differenti, diminuendo quindi la variabilità spaziale della precipitazione. Per esempio, riferendosi alla **Figura 4.15**, si considerino due bacini all'interno dello stesso distretto (evidenziati in rosso). In uno dei due piove notevolmente di più rispetto all'altro, ma nel momento in cui si effettua il *downscaling* ai nodi, ad entrambi viene attribuito lo stesso valore, perdendo così dell'informazione.



**Figura 4.15:** Rappresentazione dei dati originali di precipitazione, dei dipartimenti (in grigio scuro) e dei bacini (in grigio chiaro). Nel momento in cui si effettua il *downscaling* della precipitazione dai dipartimenti ai bacini, ai due bacini evidenziati in figura viene attribuito lo stesso valore di precipitazione (quello del dipartimento Artibonite), anche se i dati osservati di partenza sono molto diversi.

In seguito alle considerazioni appena effettuate si propone una generazione di precipitazione a partire dai nodi. Il procedimento è lo stesso del caso precedente: dai dati di precipitazione registrati dal 2002 al 2010, si generano cento anni di piogge per ogni nodo, basandosi sulle precipitazioni registrate. Dopodiché si creano cento sequenze casuali di tre anni estratte dagli anni generati, in modo da avere una previsione di precipitazione dal 2011 al 2013. Si confrontano le simulazioni ottenute immettendo come ingresso nel modello epidemiologico la precipitazione giornaliera e quella decadale.

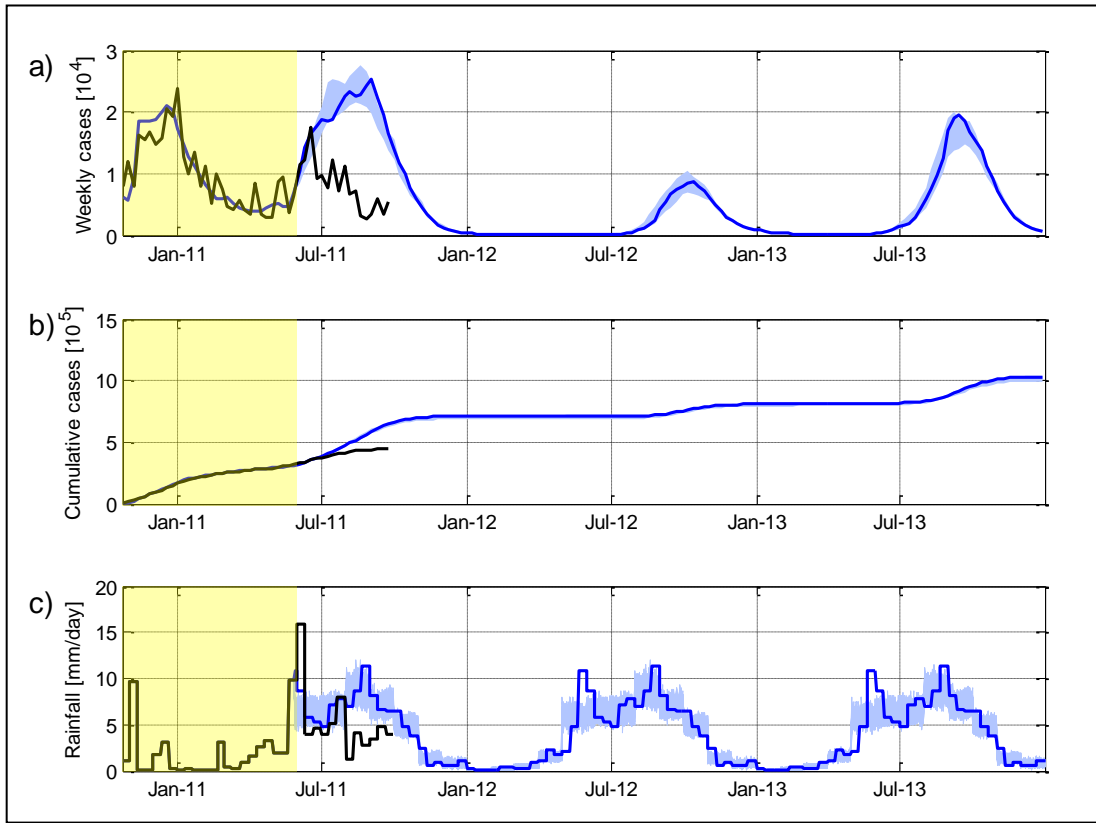
#### **4.5.1 Generazione di precipitazione partendo da dati giornalieri**

Si utilizzano le precipitazioni giornaliere fornite dall'*International Research Institute for Climate and Society*, per il periodo dal 1 gennaio 2002 al 31 dicembre 2010. In **Figura 4.16** sono riportati i risultati ottenuti con questo metodo.

Innanzitutto, nella **Figura 4.16c** si riporta, per ognuno dei cento anni di precipitazione generati, la media calcolata sui 366 bacini in cui è stata suddivisa Haiti. Si osservi, in riferimento al 2012 e al 2013, che nei mesi di maggio e giugno, la corrispondenza tra le precipitazioni generate e quelle medie non è ottima, in quanto il generatore fornisce dati più costanti rispetto alla precipitazione dell'anno medio che evidenzia, per il mese di maggio, una prima decade nella quale piove molto poco, la seconda in linea con la media mensile e la terza molto piovosa. Questo comportamento è spiegabile ricordando che il generatore lavora su base mensile e quindi non considera la variabilità intra-mensile che, invece, è confermata essere presente ad Haiti dall'andamento dell'anno di precipitazione medio. Sarebbe quindi necessario utilizzare un generatore che adotti una scala temporale più fine di quella mensile, ad esempio decadale. La predisposizione di un generatore con parametri a scala decadale richiede però tempi di calcolo che vanno al di là degli scopi di questo lavoro, che guarda alla definizione di scenari di massima di risorgenza della malattia, piuttosto che alla ricostruzione puntuale di possibili *output* epidemiologici. Tuttavia, la definizione di un generatore di precipitazione più preciso rappresenta una condizione necessaria all'utilizzo del modello a scopo predittivo, in

un orizzonte temporale di qualche mese, invece che, come in questo caso, di qualche anno. Inoltre, per generare insiemi di scenari più vari possibile, si potrebbe anche introdurre una variabilità inter-annuale, che vada a modificare il regime pluviometrico a livello stagionale, in quanto il modello epidemiologico si dimostra molto sensibile a queste variazioni. Infatti, la stagione particolarmente secca dell'estate 2011 è probabilmente responsabile del più rapido declino dei casi osservati rispetto alle simulazioni effettuate.

In **Figura 4.16a** gli scenari generati con la precipitazione originata dal generatore non si discostano molto da quelli dell'anno di precipitazione medio. Infatti, le simulazioni con la precipitazione generata prospettano un numero di casi cumulati che a fine 2013 sono compresi tra 994 870 e 1 036 750, mentre utilizzando l'anno di precipitazione medio sono previsti 1 030 260 casi cumulati di colera (**Figura 4.16b**). Utilizzando questo metodo, quindi, la previsione non si discosta molto da quella effettuata con l'anno medio. Questo fatto si spiega considerando che la generazione della precipitazione avviene indipendentemente per ogni. Per simulare situazioni di siccità in tutta l'isola o di estrema piovosità sarebbe necessario generare molti più anni di precipitazione rispetto ai cento attualmente generati, ma ne conseguirebbe un carico computazionale molto oneroso. Una possibile soluzione potrebbe essere quella di introdurre dei raggruppamenti dei bacini, in modo da imporre un regime precipitativo simile per alcune aree (come ad esempio è stato fatto per i dipartimenti, **Capitolo 4.4.2**) e quindi imponendo una forte correlazione spaziale della precipitazione, che comporta la generazione di condizioni più estreme con conseguenti scenari epidemiologici più o meno gravi.



**Figura 4.16:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei nodi.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei nodi.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In nero sono raffigurati i dati reali, mentre in blu sono indicate le precipitazioni medie. La fascia azzurra rappresenta il *range* delle precipitazioni giornaliere (mediate sui nodi per ognuno dei cento anni generati) generate con il generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei nodi.

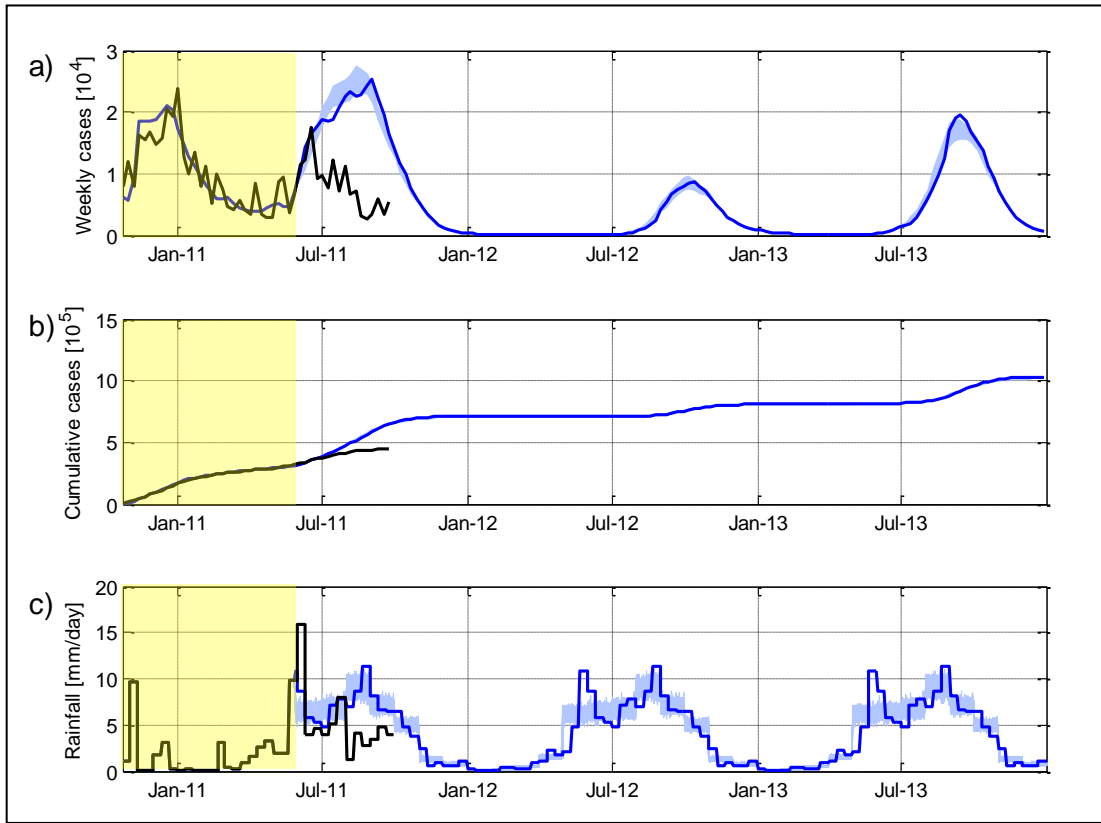
Lo sfondo giallo indica il periodo della calibrazione.



## 4.5.2 Generazione di precipitazione partendo da dati decadali

Si immettono come ingresso nel generatore puntuale di precipitazione di Poisson, le precipitazioni decadali fornite dall'*International Research Institute for Climate and Society*, per il periodo dal 1 gennaio 2002 al 31 dicembre 2010.

Osservando la **Figura 4.17c**, che riporta le medie sui nodi delle precipitazioni giornaliere, è possibile notare che la precipitazione generata devia in modo poco sensibile dalla media mensile e non presenta variazioni intra-mensili, che per essere colte richiederebbero l'adozione di una scala temporale più fine. Anche i casi cumulati previsti per la fine del 2013 con la precipitazione generata da dati decadali sono simili a quelli previsti dalla precipitazione media (**Figura 4.17b**). Precisamente le simulazioni eseguite con la precipitazione generata da dati decadali prospettano un numero di casi cumulati compreso tra 1010 990 e 1037 870 (mentre nel caso si utilizzi la precipitazione media sono di 1030 260). Il limitato spettro di scenari prodotto (**Figura 4.17a**) è dovuto al fatto che, così come nel caso precedente, la generazione avviene indipendentemente per ogni nodo e quindi con soli cento anni di precipitazione generati non si riescono a riprodurre situazioni in cui vaste aree hanno una precipitazione o molto scarsa o molto abbondante.



**Figura 4.17:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione decadale dei nodi.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati e in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dalla precipitazione decadale dei nodi.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In nero sono raffigurati i dati reali, mentre in blu sono indicate le precipitazioni medie. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi della precipitazione giornaliera (mediata sui nodi per ognuno dei cento anni generati) generata con il generatore puntuale di Poisson a partire dalla precipitazione decadale dei nodi.

Lo sfondo giallo indica il periodo della calibrazione.

## 4.6 Simulazione con la precipitazione calcolata con modelli ARX

Tutti i metodi precedenti si basano su statistiche della serie storica di precipitazione. In questo paragrafo si propone invece un modello di previsione delle precipitazioni che, partendo dall'analisi della serie storica di precipitazioni a disposizione, genera dei dati di pioggia considerando l'andamento pluviometrico dell'anno in corso e quindi sfrutta al massimo le informazioni più recenti. Inoltre, si vorrebbe considerare anche la correlazione spaziale delle precipitazioni, in quanto la precipitazione di zone confinanti è quasi sempre correlata (**Capitolo 3.5**). Si vuole quindi introdurre un modello di previsione di precipitazione che consideri sia la correlazione temporale che quella spaziale.

Per il caso di studio si propone l'utilizzo di un modello ARX (*Auto Regressive Exogenous*), in cui la correlazione temporale è descritta dalla parte auto-regressiva del modello, mentre quella spaziale dall'ingresso esogeno. Si considera separatamente la generazione della precipitazione relativa ad ogni distretto. La parte auto-regressiva corrisponde alla precipitazione di un dato dipartimento registrata nei giorni precedenti, mentre l'ingresso esogeno corrisponde alle precipitazioni degli altri distretti.

Un modello lineare autoregressivo con ingresso esogeno (ARX) è definito come:

$$y_t = \sum_{i=1}^{n_a} a_i y_{t-i} + \sum_{i=1}^{n_b} b_i u_{t-i} + e_t \quad (4.1)$$

dove  $y_t$  è la previsione di precipitazione che si vuole effettuare,  $y_{t-i}$  è la pioggia ai tempi precedenti del distretto preso in esame,  $u_{t-i}$  è la precipitazione dei distretti circostanti ai tempi precedenti ed  $e_t$  è un rumore bianco. Nel caso in esame  $u_{t-i}$  sarà un vettore con nove elementi, corrispondente al numero dei distretti meno quello in esame.

Per la calibrazione dei coefficienti sono state utilizzate le precipitazioni decadalì fornite dall'*International Research Institute for Climate and Society*, dal 1 gennaio 2002 al 31 maggio 2011.

Non è scontato che la precipitazione di un distretto sia influenzata in egual modo da quella di tutti gli altri distretti, infatti si suppone che ci sia una maggiore correlazione tra dipartimenti confinanti. Si è scelto quindi di non utilizzare lo stesso ordine  $n_b$  per tutti gli ingressi esogeni, ma di calibrarlo opportunamente per ogni dipartimento, imponendo che sia l'ordine di  $n_a$  che quello di  $n_b$  siano al massimo uguali a 3. Questo equivale all'assunzione che la precipitazione al tempo  $t$  sia influenzata al più dalla precipitazione caduta durante le antecedenti tre decadi.

Per scegliere la complessità del modello ARX, ovvero stabilire l'ordine per la regressione, sia della precipitazione del distretto che per quella degli ingressi esogeni, sono stati calcolati gli AIC (*Akaike Information Criterion*) relativi a modelli con differenti combinazioni degli ordini. Si ricorda che l'AIC è dato dalla somma di due addendi, il primo dei quali penalizza la complessità del modello, mentre il secondo misura sinteticamente l'aderenza del modello ottimo di ordine  $n$  ai dati usati per l'identificazione, ed è quindi decrescente con l'ordine  $n$  [Bittanti, 2000].

Dopo aver calibrato indipendentemente ogni modello ARX con i dati di precipitazione decadale, è stato calcolato il valore dell'AIC come:

$$AIC = 2\Theta + \eta \ln\left(\frac{RSS}{\eta}\right)$$

dove  $\Theta$  è il numero di parametri del modello,  $\eta$  è il numero dei dati (in particolare  $\eta = n_d n_w$  con  $n_d$  numero di dipartimenti e  $n_w$  numero delle settimane di registrazione dei dati), mentre  $RSS$  (*Residual Sum of Square*), che è una misura della discrepanza tra il modello preso in considerazione e i dati, è calcolabile come:

$$RSS = \sum_{i=1}^{n_d} \sum_{j=1}^{n_w} [P(i, j) - \hat{P}(i, j)]^2$$

con  $P(i, j)$  e  $\hat{P}(i, j)$  che sono rispettivamente la precipitazione decadale reale e la precipitazione decadale simulata dal modello per un dato distretto. Tramite questo criterio è possibile scegliere, tra un insieme di modelli, quello che fornisce la prestazione migliore per l'insieme dei dati in esame, in termini di aderenza alla precipitazione osservata e di parsimonia del modello. In particolare, in questo caso, sono stati selezionati i tre modelli migliori per ogni distretto e di conseguenza si avranno tre diverse previsioni di precipitazione. Nella **Tabella 4.1** sono riportati, per i tre modelli migliori per ogni distretto, il valore dell'AIC che serve per confrontare i diversi modelli, l'ordine di autoregressione e l'ordine di autoregressione degli ingressi esogeni. È possibile notare che gli AIC dei modelli relativi a un dipartimento sono molto simili tra loro. Per quanto riguarda l'ordine di autoregressione e quello dell'ingresso esogeno, l'importanza dei singoli dipartimenti dipende dal valore assunto dai coefficienti. Per meglio chiarire il procedimento utilizzato, a titolo di esempio si riportano i coefficienti del primo caso, quindi relativi al primo modello del distretto di Nord-Ovest (prima riga della **Tabella 4.1**).

Indicando con  $z^{-1}$  l'operatore di ritardo unitario, ovvero:

$$z^{-1}y(t) = y(t - 1); z^{-2}y(t) = y(t - 2); \dots; z^{-i}y(t) = y(t - i)$$

ed introducendo i polinomi:

$$\begin{aligned} A(z) &= 1 - a_1z^{-1} - a_2z^{-2} - \dots - a_{n_a}z^{-n_a} \\ B(z) &= b_1 - b_2z^{-1} - b_3z^{-2} - \dots - b_{n_b}z^{-n_b+1} \end{aligned}$$

la Formula (4.1) può anche essere scritta come:

$$A(z)y(t) = B(z)u(t - 1) + e(t)$$

Distretto	AIC	$n_a$	$u_1$	$u_2$	$u_3$	$u_4$	$u_5$	$u_6$	$u_7$	$u_8$	$u_9$
	1.88478	1	1	1	1	3	1	1	1	1	2
<b>Nord-Ovest</b>	1.88554	1	3	1	1	1	1	1	2	1	2
	1.88560	1	1	1	1	3	1	1	2	1	2
	3.30511	1	1	1	2	3	1	1	3	1	1
<b>Nord</b>	3.30559	1	1	1	2	3	2	3	1	1	1
	3.30560	1	1	1	2	3	2	1	1	2	1
	3.76390	1	1	1	1	3	1	1	1	1	1
<b>Nord-Est</b>	3.76421	1	1	1	1	3	1	3	1	1	1
	3.76438	1	3	1	1	3	1	1	1	1	1
	3.40337	2	3	1	1	3	1	1	3	1	3
<b>Artibonite</b>	3.40436	2	3	1	1	3	1	1	1	1	2
	3.40670	2	3	1	1	3	1	1	3	1	1
	4.00123	3	3	1	1	2	1	1	1	1	1
<b>Centre</b>	4.00364	3	3	2	2	2	1	1	1	1	1
	4.00574	3	3	1	1	3	1	1	1	1	1
	2.89609	1	1	1	2	3	3	3	1	3	1
<b>Grande Anse</b>	2.89657	2	1	1	2	3	3	3	1	3	1
	2.89854	2	1	1	2	2	3	3	1	3	1
	3.12930	3	1	3	1	2	1	1	2	3	2
<b>Nippes</b>	3.13148	3	1	3	1	2	1	1	3	3	2
	3.13274	3	1	3	1	1	1	1	2	3	2
	3.14770	1	1	1	2	3	3	1	3	3	1
<b>Ovest</b>	3.14879	1	1	1	1	3	3	1	3	3	1
	3.14960	1	1	1	2	2	3	1	3	3	1
	3.11813	3	1	2	1	3	1	1	3	2	2
<b>Sud</b>	3.11973	3	1	3	1	3	1	1	3	2	2
	3.12168	3	1	1	3	3	1	1	3	2	2
	3.02340	2	1	2	1	1	3	1	3	2	3
<b>Sud-Est</b>	3.02452	2	1	1	2	2	3	1	3	2	3
	3.02538	2	1	1	2	1	3	1	3	2	3

**Tabella 4.1:** La tabella riporta i tre modelli ARX migliori per ogni distretto. La prima colonna corrisponde ai distretti, per ognuno dei quali sono riportati i valori relativi ai tre modelli migliori secondo il criterio AIC. La terza colonna riporta l'ordine di autoregressione, mentre dalla quarta all'ultima colonna sono indicati gli ordini di regressione per i diversi ingressi esogeni.

Dove:

$$A = [1 \quad -0.0515] \quad B = \begin{bmatrix} -0.1057 & 0 & 0 \\ 0.0585 & 0 & 0 \\ -0.0556 & 0 & 0 \\ 0.0963 & 0.12709 & 0.0652 \\ -0.0705 & 0 & 0 \\ -0.1434 & 0 & 0 \\ 0.0017 & 0 & 0 \\ 0.1867 & 0 & 0 \\ 0.0987 & -0.1140 & 0 \end{bmatrix} \quad e(t) = 6.0987$$

$$y(t) = \begin{bmatrix} y(t) \\ y(t-1) \end{bmatrix}$$

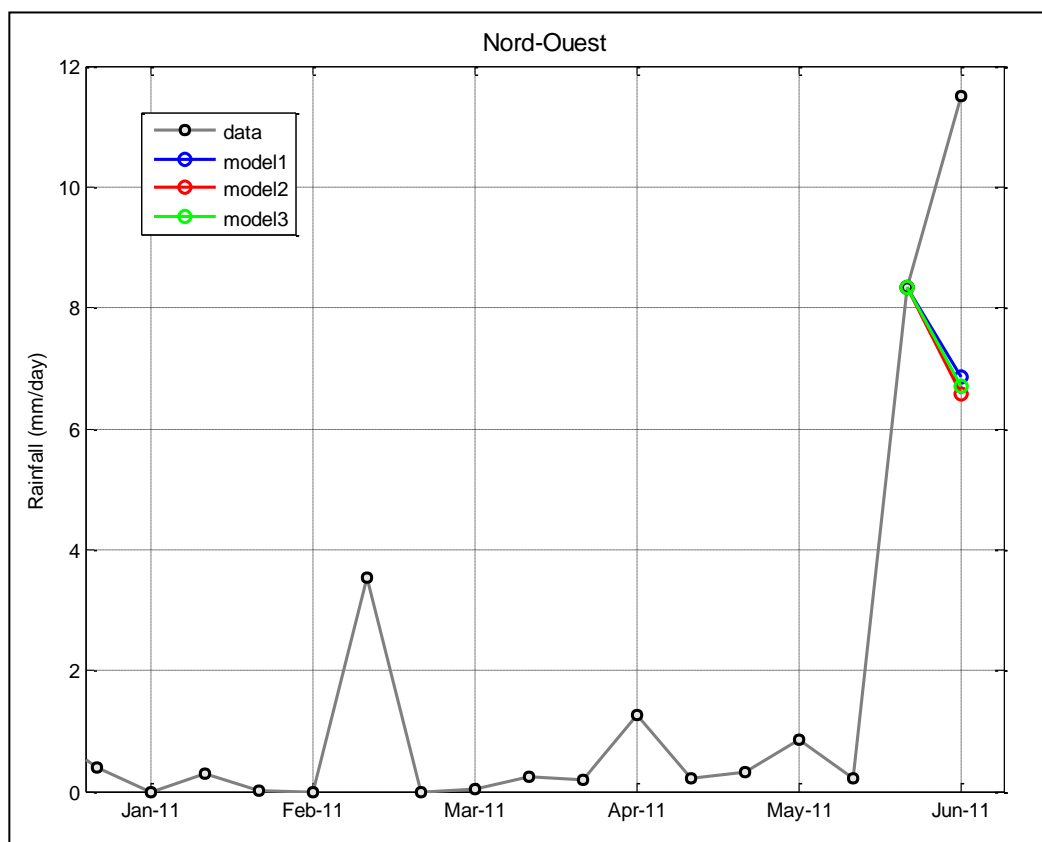
$$u(t-1) =$$

$$\begin{bmatrix} u_1(t-1) & u_2(t-1) & u_3(t-1) & u_4(t-1) & u_5(t-1) & u_6(t-1) & u_7(t-1) & u_8(t-1) & u_9(t-1) \\ u_1(t-2) & u_2(t-2) & u_3(t-2) & u_4(t-2) & u_5(t-2) & u_6(t-2) & u_7(t-2) & u_8(t-2) & u_9(t-2) \\ u_1(t-3) & u_2(t-3) & u_3(t-3) & u_4(t-3) & u_5(t-3) & u_6(t-3) & u_7(t-3) & u_8(t-3) & u_9(t-3) \end{bmatrix}$$

La matrice A è di dimensioni (1x2), dove il numero delle colonne dipende dal valore dell'ordine  $n_a$ , in particolare sarebbe  $n_a+1$ . Invece B è una (9x3), dove 9 corrisponde al numero degli ingressi esogeni, ovvero alla precipitazione degli altri distretti, e 3 è l'ordine massimo  $n_b$ . I coefficienti sono pari a zero dove l'ordine della regressione è minore di 3.

Con questo metodo è possibile compiere delle previsioni di precipitazione per un dato distretto, utilizzando la precipitazione del distretto stesso e quella relativa agli altri distretti ai tempi precedenti. Attraverso un modello ARX è però possibile predire la precipitazione solo per la prima decade di simulazione (nel caso di studio dall'1 al 10 giugno 2011), poiché si dimostra che la precipitazione prevista a più di un passo di previsione risulta del tutto scorrelata dall'ingresso del modello.

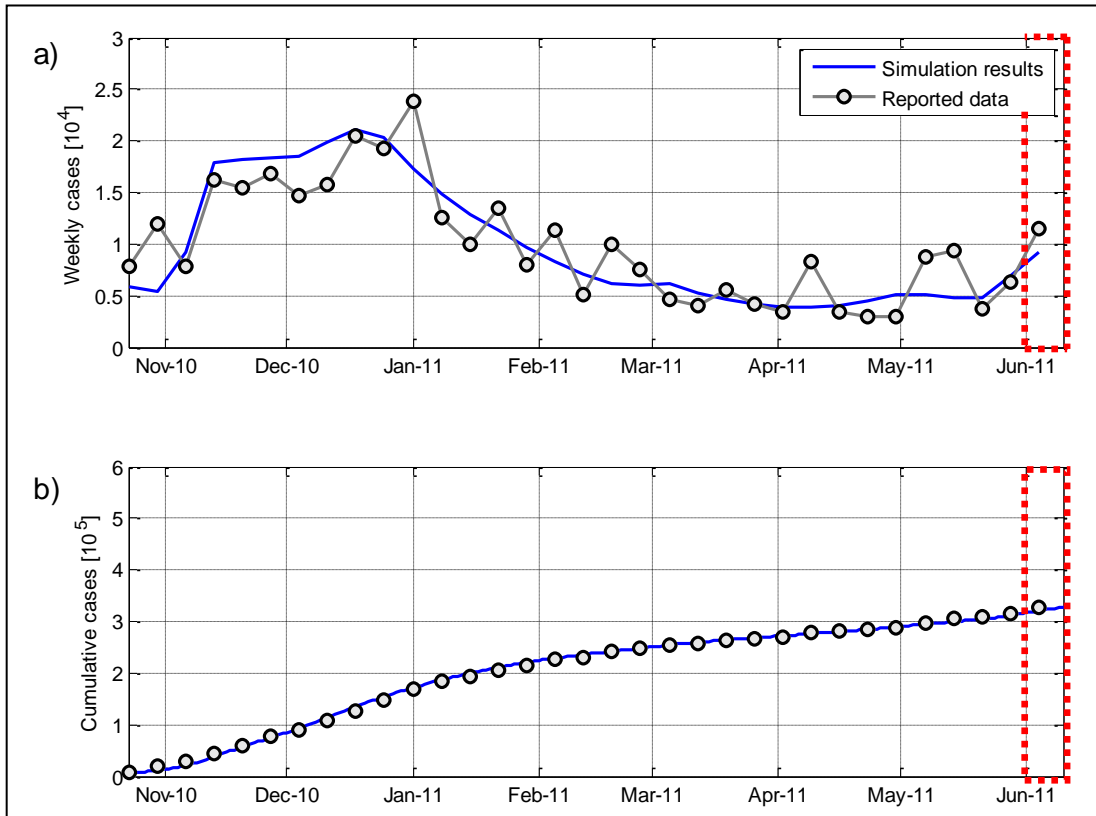
Per il distretto Nord-Ovest si riporta la previsione di precipitazione effettuata, utilizzando i tre modelli ARX selezionati (**Figura 4.18**). Com'è possibile notare i tre modelli restituiscono una previsione molto simile. Dopo aver previsto una decade di precipitazione per ogni distretto, si utilizza il modello SIRB con le precipitazioni (**Capitolo 2, Modello 2.4**) per effettuare una previsione dei casi di colera per la prima settimana di giugno.



**Figura 4.18:** Precipitazione simulata con i tre migliori modelli ARX per il distretto Nord-Ovest. Si noti che ogni dato rappresenta quanto piove in un giorno di una data decade. La previsione si riferisce alla prima decade di giugno 2011.

Nella **Figura 4.19** viene riportata la previsione dell'andamento dell'epidemia realizzabile solo per l'ultimo dato, quello relativo alla prima settimana di giugno. Come ingresso di precipitazione si utilizzano per il periodo dal 23 ottobre 2010 al 31 maggio 2011 le precipitazioni decadali della *NASA-JAXA's Tropical Rainfall Measuring Mission* e dal 1 giugno 2011 al 10 giugno 2011 la precipitazione prevista con il modello ARX sulla base delle precipitazioni decadali fornite dall'*International Research Institute for Climate and Society*. Poiché i tre modelli di precipitazione relativi a un dipartimento restituiscono una precipitazione molto simile, nel momento in cui si inseriscono come ingresso nel modello epidemiologico si ottiene lo stesso numero di casi di colera previsti, quindi il grafico di **Figura 4.19** riporta solo una previsione.

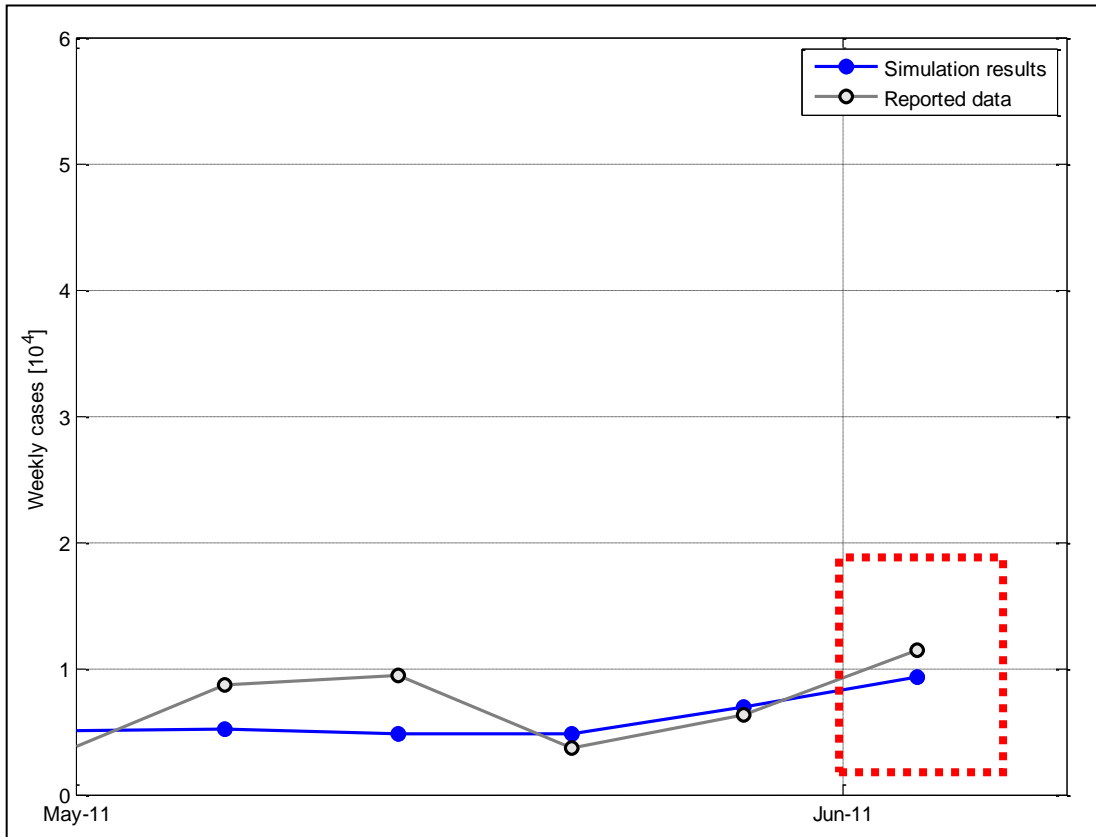




**Figura 4.19:** Simulazione dei casi di colera per tutta Haiti eseguita con il modello epidemiologico (Capitolo 2, Modello 2.4) utilizzando i modelli ARX per prevedere le precipitazioni di inizio giugno 2011. Il grafico (a) rappresenta i casi settimanali di colera, mentre il grafico (b) quelli cumulati. In blu sono riportati i dati simulati e con i pallini grigi i dati reali. Il dato dei casi di colera simulato con la precipitazione generata dai modelli ARX è contornato con un tratteggio rosso.

Si osservi che mentre per la precipitazione l'intervallo temporale erano le decadi, per i casi di colera i dati hanno cadenza settimanale. Di conseguenza, se le precipitazioni sono previste fino al 10 giugno, è possibile simulare la diffusione dell'epidemia di colera solo fino al 4 giugno (una settimana dopo il 28 maggio).

Con questo metodo di previsione della precipitazione, i casi di colera settimanali previsti per la prima settimana di giugno sono 9 264, mentre con il modello che utilizza le precipitazioni medie si calcolano 9 016 casi di colera, considerando che nella realtà sono stati registrati 11 450 casi (**Figura 4.20**). Il vantaggio di utilizzare un modello ARX per generare la precipitazione risiede nel fatto che con questo metodo si sfruttano le informazioni relative all'andamento pluviometrico in corso.



**Figura 4.20:** Zoom temporale sulla simulazione dei casi di colera simulati con il modello epidemiologico (Capitolo 2, Modello 2.4) utilizzando i modelli ARX per prevedere le precipitazioni di inizio giugno 2011. In blu sono riportati i dati generati e con i pallini grigi i dati reali. Il dato dei casi di colera simulato con la precipitazione generata dai modelli ARX, che si riferisce al 4 giugno 2011, è contornato con un tratteggio rosso.

La previsione di precipitazione con il modello ARX comporta quindi un miglioramento nella previsione a breve termine, ma non permette di effettuare previsioni per orizzonti temporali maggiori di una decade. Il modello ARX può però essere impiegato in una politica di controllo in linea dell'epidemia, in cui le previsioni a dieci giorni permettono di implementare degli interventi su brevissimo periodo, ad esempio inviando forniture di acqua potabile o kit igienico-sanitari nelle zone che si prevede saranno maggiormente colpite.

## **4.7 Confronto tra gli *output* epidemiologici ottenuti con i diversi metodi di previsione della precipitazione**

In generale è possibile affermare che il modello è molto sensibile all'ingresso di precipitazione, in quanto il numero di casi di colera aumenta con l'aumentare dell'intensità e delle fluttuazioni della precipitazione. Infatti, come per il caso dell'anno tipico, oscillazioni maggiori comportano risposte più impulsive nella diffusione della malattia. Per quanto riguarda la previsione a lungo termine, tutte le simulazioni effettuate concordano nel prevedere una prima risorgenza dell'epidemia intorno a settembre 2012 e una seconda più acuta a settembre 2013.

In **Tabella 4.2** si riporta uno schema riassuntivo dei casi cumulati previsti per il 30 settembre 2011. Per ogni metodo sono anche mostrati la differenza con i casi reali e gli errori di stima.

Nonostante sia il metodo più semplice, tra i diversi metodi di previsione della precipitazione analizzati, quello che restituisce una maggiore variabilità degli scenari è il metodo che si basa sulle statistiche di precipitazione (media, 25-esimo percentile e 75-esimo percentile). Infatti, in questo caso si riescono a cogliere le situazioni estreme, in quanto si impone che su tutta Haiti la precipitazione sia particolarmente scarsa (25-esimo percentile) oppure particolarmente abbondante (75-esimo percentile). In particolare, la simulazione che meglio riesce a riprodurre il reale andamento dell'epidemia nel terzo trimestre del 2011 è quella effettuata con il 25-esimo percentile della precipitazione. Il motivo per il quale l'epidemia è meglio descritta dal 25-esimo percentile della precipitazione, piuttosto che dalla simulazione effettuata con l'anno di precipitazione medio, è che il 2011 risulta essere un anno particolarmente secco (quindi più vicino al 25-esimo percentile che alla media), in cui la stagione piovosa si è verificata solo nel periodo limitato tra maggio e luglio, quando in genere si protrae per un intervallo più lungo, che va da maggio fino a ottobre.

			Casi cumulati previsti (casi)	Differenza rispetto ai casi reali (casi)	Errore 1 <sup>9</sup> (%)	Errore 2 <sup>10</sup>	
<b>CASI CUMULATI REALI REGISTRATI (al 30 settembre 2011)</b>			452 710	-	-	-	
<b>Simulazione con precipitazione reale</b>			518 000	65 290	14	2.86E+10	
<b>Simulazione con statistiche di precipitazione</b>	media		647 220	194 510	43	5.14E+10	
	25-esimo percentile		485 000	32 290	7	2.84E+10	
	75-esimo percentile		772 000	319 290	71	1.12E+11	
<b>Simulazione con precipitazione dell'anno tipico</b>			667 440	214 730	47	5.09E+10	
<b>Simulazione con il generatore puntuale di precipitazione di Poisson</b>	distretti	distretti disaggregati	caso MIGLIORE	582 980	130 270	29	3.47E+10
			caso PEGGIORE	692 310	239 600	53	7.66E+10
		distretti aggregati	caso MIGLIORE	572 990	120 280	27	3.39E+10
			caso PEGGIORE	684 810	232 100	51	6.78E+10
	nodi	precipitazioni giornaliere	caso MIGLIORE	624 060	171 350	38	4.50E+10
			caso PEGGIORE	657 030	204 320	45	5.81E+10
		precipitazioni decadali	caso MIGLIORE	633 560	180 850	40	4.82E+10
			caso PEGGIORE	660 240	207 530	46	5.83E+10

**Tabella 4.2:** Per ogni metodo di previsione di precipitazione è stato riportato il numero di casi di colera cumulati previsti dal modello epidemiologico per il 30 settembre 2011, la differenza rispetto a quelli reali e l'errore di stima commesso, calcolato sia rispetto al totale dei casi cumulati alla data di riferimento che considerando gli scarti tra i casi a tutti i passi temporali.

<sup>9</sup> Calcolato come:  $Errore\ 1 = \frac{casi\ stimati\ (al\ 30\ settembre\ 2011) - casi\ reali\ (al\ 30\ settembre\ 2011)}{casi\ reali\ (al\ 30\ settembre\ 2011)} \cdot 100$

<sup>10</sup> Calcolato come:  $Errore\ 2 = \sum (casi\ stimati - casi\ osservati)^2$ , calcolati per ogni nodo e per tutti i passi temporali.

## CONCLUSIONI

Per poter implementare delle politiche di controllo e previsione dell'epidemia di colera che sta colpendo Haiti da ottobre 2010, si devono realizzare delle simulazioni, su breve e lungo periodo, sull'evoluzione di tale malattia. Nella presente tesi si è constatato che introducendo la precipitazione come variabile nel modello epidemiologico si riesce a descrivere meglio il reale andamento dell'epidemia di colera. Inoltre, inserendo uno spettro di previsioni di precipitazioni è possibile generare un vasto numero di scenari, che possono servire da supporto per pianificare i possibili interventi da attuare per arginare la diffusione dell'epidemia.

Per stabilire il metodo più appropriato di generazione della precipitazione, sono state vagliate diverse alternative:

- Media, 25-esimo percentile e 75-esimo percentile della precipitazione
- Precipitazioni dell'anno tipico
- Metodo di previsione con modelli ARX
- Metodi di generazione puntuale della precipitazione di tipo poissoniano partendo dai dipartimenti e dai nodi.

Il modello ha mostrato buone prestazioni nel periodo di validazione (dal 29 maggio al 30 settembre 2011), per il quale è stata utilizzata in ingresso la precipitazione effettivamente osservata.

In generale, è opportuno compiere delle previsioni a lungo termine da integrare con quelle a breve termine. In questo modo si possono pianificare dei programmi di intervento nel lungo periodo, che possono poi essere modificati in base alle esigenze che si vengono via via a verificare, con degli interventi di brevissimo periodo.

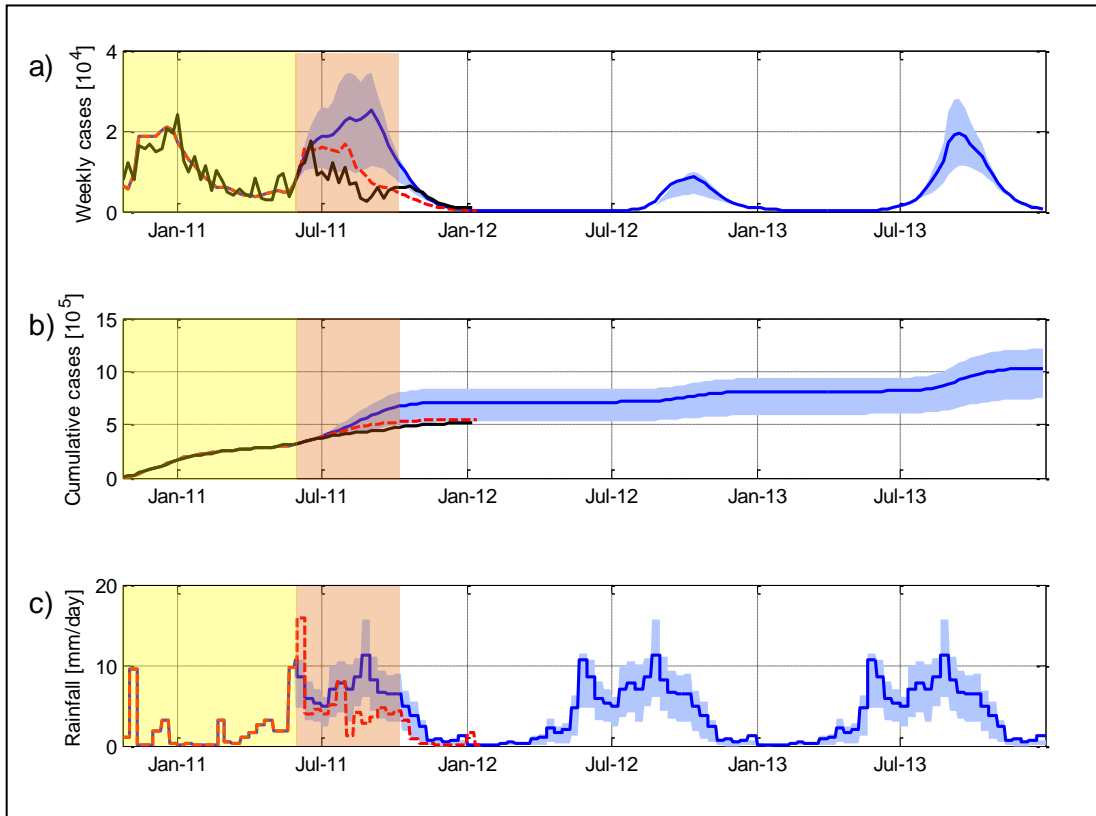
Le previsioni a lungo termine sono necessarie per pianificare strategie impegnative, sia da un punto di vista temporale che economico. Fanno parte di questa categoria interventi atti a costruire infrastrutture idriche adeguate, a migliorare le condizioni igienico-sanitarie e a sviluppare programmi di sensibilizzazione della popolazione. Il gran numero di scenari prodotto consente di valutare la linea di azione migliore. Tutte le simulazioni concordano nel prevedere una prima risorgenza dell'epidemia di colera ad Haiti intorno a settembre 2012 e una seconda, più acuta, a settembre 2013. Si nota quindi l'andamento stagionale dell'epidemia che segue i periodi piovosi. Il metodo migliore per generare la precipitazione da introdurre nel modello epidemiologico è quello che utilizza le statistiche di precipitazione (media, 25-esimo percentile e 75-esimo percentile). Infatti, nonostante sia il metodo più semplice è comunque da preferire rispetto agli altri poiché offre una maggiore variabilità di scenari. Per generare un più ampio spettro di scenari anche con gli altri metodi, si dovrebbero apportare alcune migliorie, per esempio sarebbe opportuno raffinare la scala temporale del generatore puntuale di precipitazione poissoniano, oppure introdurre un termine per considerare esplicitamente la correlazione spaziale dei bacini.

Per quanto riguarda le previsioni a breve termine, un modello robusto di generazione della precipitazione è il modello ARX, che considera esplicitamente le correlazioni sia temporali che spaziali. L'utilizzo, da parte dei modelli ARX, di tutte le informazioni disponibili li rende inoltre particolarmente adatti all'utilizzo in una politica di controllo in linea dell'epidemia. Alcuni interventi che possono essere implementati anche a brevissimo termine sono, ad esempio, l'approvvigionamento di

acqua potabile o la spedizione di kit igienico-sanitari nelle zone che si prevede saranno più colpite.

Infine, si effettua un confronto tra gli scenari generati con la precipitazione media e i dati dei casi di colera aggiornati al 20 gennaio 2012 (**Figura 5.1**). L'andamento delle precipitazioni reali registrate dopo il 30 settembre 2011 segue quello dell'anno di precipitazione medio (**Figura 5.1c**), diminuendo in corrispondenza dell'inizio della stagione secca (ultima decade di ottobre 2011). Anche i casi settimanali di colera osservati diminuiscono (**Figura 5.1a**), in accordo con le simulazioni effettuate e quindi la differenza tra i casi di colera cumulati reali e simulati diminuisce (**Figura 5.1b**). Pertanto le previsioni eseguite riescono a riprodurre il reale andamento dell'epidemia di colera a meno di situazioni particolari, come può essere considerata la mancanza della seconda stagione piovosa durante l'estate 2011.

Effettuando delle previsioni affidabili sulla diffusione del colera e definendo un gran numero di possibili scenari di precipitazione e di conseguenza anche epidemiologici, i modelli analizzati in questa tesi, rappresentano un efficace strumento di supporto decisionale durante tutta l'evoluzione dell'epidemia.



**Figura 5.1:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con l'anno di precipitazione medio. Confronto con i dati reali aggiornati al 20 gennaio 2012.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, con una linea tratteggiata rossa quelli simulati utilizzando i dati di precipitazione reali e in blu quelli che utilizzano la precipitazione mediata sui dieci anni. Infine, la banda azzurra rappresenta l'intervallo dei casi simulati utilizzando il 25-esimo percentile e 75-esimo percentile delle precipitazioni per ogni nodo.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni. In rosso è rappresentato l'andamento della pioggia decadale secondo i dati della *NASA-JAXA's Tropical Rainfall Measuring Mission*, mentre in blu quello delle precipitazioni medie, usate per la simulazione. La banda azzurra rappresenta l'intervallo tra il primo quartile e il terzo quartile delle precipitazioni.

Lo sfondo giallo indica il periodo della calibrazione, quello rosa indica il periodo dal 29 maggio 2011 al 30 settembre 2011.



## 5.1 Sviluppi futuri

Nella presente tesi sono stati analizzati diversi metodi per prevedere le precipitazioni da inserire come ingresso nel modello di evoluzione del colera. Poiché il modello epidemiologico si è dimostrato molto sensibile a questo ingresso, si potrebbero apportare delle modifiche al generatore puntuale di precipitazione di Poisson, oppure si potrebbero utilizzare altri metodi di generazione delle precipitazioni per verificare se è possibile affinare le previsioni sull'andamento dell'epidemia.

In primo luogo, si potrebbe utilizzare il generatore puntuale di precipitazione di Poisson generando un maggior numero di anni di pioggia, in modo da produrre più scenari possibili (il che però renderebbe il calcolo molto più oneroso dal punto di vista computazionale). In alternativa, si potrebbe modificare il generatore originando la precipitazione non più basandosi sulle medie mensili, ma su quelle decadali. Infatti, così facendo, si considererebbe anche la variabilità intra-mensile, che con l'attuale generatore viene eliminata. Nella generazione della precipitazione a partire dai nodi si potrebbe invece inserire un termine per considerare esplicitamente la correlazione spaziale della precipitazione tra bacini vicini.

Si potrebbero inoltre utilizzare altri tipi di generatori, per esempio un generatore puntuale di precipitazione *rectangular pulse*, che considera tra le variabili anche la durata dell'evento precipitativo, o altri modelli più complessi come quello di Bartlett-Lewis o di Neymann-Scott.

Infine, sarebbe molto interessante generare la precipitazione con un generatore che consideri anche la struttura spaziale del territorio, come ad esempio quello proposto da *Brisette et al.* (2007) che introduce la matrice di correlazione spaziale nella generazione.

## BIBLIOGRAFIA

Andrews J. & S. Basu (2011). *Transmission dynamics and control of cholera in Haiti: an epidemic model*. *Lancet*, 377, 1248-1252.

Bertuzzo E., M. Gatto, I. Rodriguez-Iturbe & A. Rinaldo (2007). *River networks and ecological corridors: Reactive transport on fractals, migration fronts, hydrochory*. *Water Resources Research*, 43, W04419.

Bertuzzo E., S. Azaele, A. Maritan, M. Gatto, I. Rodriguez-Iturbe & A. Rinaldo (2008). *On the space-time evolution of a cholera epidemic*. *Water Resources Research*, 44, W01424.

Bertuzzo E., R. Casagrandi, I. Rodriguez-Iturbe & A. Rinaldo (2010). *On spatially explicit models of cholera epidemics*. *Journal of The Royal Society Interface*, 7, 321-333.

Bertuzzo E., L. Mari, L. Righetto, M. Gatto, R. Casagrandi, M. Blokesch, I. Rodriguez-Iturbe & A. Rinaldo (2011). *Prediction of the spatial evolution and effects*

*of control measures for the unfolding Haiti cholera outbreak.* Geophysical Research Letters, 38, L06403.

Biermann J. (2011). *Hydrologic and Anthropogenic Controls of Cholera Transmission: A Case Study of Haiti Epidemic.* Master Thesis. Laboratory of Ecohydrology, École Polytechnique Fédérale de Lausanne.

Bittanti S. (2000). *Identificazione dei Modelli e Controllo Adattivo.* Pitagora.

Brisette F. P., M. Khalili & R. Leconte (2007). *Efficient stochastic generation of multi-site synthetic precipitation data,* Journal of Hydrology 345, 121-133.

Capasso V. & S. Paveri-Fontana (1979). *A mathematical model for the 1973 cholera epidemic in the European Mediterranean region.* Rev Epidémiol Santé Publique 27, 121–132.

Capasso V. (1993). *Mathematical structures of epidemic systems.* Berlin, D: Springer-Verlag.

CDC (2011). *2010 Haiti Cholera Outbreak Maps - Map of Cumulative Attack Rate With Data on Cases, Hospitalizations, and Deaths.* Centers for Disease Control and Prevention. <[http://www.bt.cdc.gov/situationawareness/haiticholera/map\\_1.asp](http://www.bt.cdc.gov/situationawareness/haiticholera/map_1.asp)>

Chao D., M. Halloran & I. Longini (2011). *Vaccination strategies for epidemic cholera in Haiti with implications for the developing world.* Proceedings of the National Academy of the United States of America, 108, 7081-7085.

CIA (2009). *2009 CIA's World Factbook.* [On line] Disponibile a: <<https://www.cia.gov/library/publications/the-world-factbook/index.html>>.

CIA (2010). *2010 CIA's World Factbook.* [On line] Disponibile a: <<https://www.cia.gov/library/publications/the-world-factbook/geos/ha.html>>.

Codeço C. T. (2001). *Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir*. BMC Infectious Diseases, 1, 1.

Colwell R. (1996). *Global climate and infectious disease: the cholera paradigm*. Science, 274, 2025-2031.

Comstock L. E., D. Maneval Jr, P. Panigrahi, A. Joseph, M. M. Levine, J. B. Kaper, J. G. Morris Jr & J. A. Johnson (1995). *The capsule and O antigen in Vibrio cholerae O139 Bengal are associated with a genetic region not present in Vibrio cholerae O1*. Infection and Immunity, 63, 317-323.

Eggo C. S., R. M. & N. M. Ferguson (2010). *Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States*. Jurnal of The Royal Society Interface.

Hartley D.M., J. G. Morris & D. L. Smith (2006). *Hyperinfectivity: a critical element in the ability of Vibrio cholerae to cause epidemics?*. PLOS Medicine, 3, 63-69.

Hendriksen R. S., L. B. Price, J. M. Schupp, J. D. Gillece, R. S. Kaas, D. M. Engelthaler, V. Bortolaia, T. Pearson, A. E. Waters, B. P. Upadhyay, S. D. Shrestha, S. Adhikari, G. Shakya, P. S. Keim & F. M. Aarestrup (2011). *Population Genetics of Vibrio Cholerae from Nepal in 2010: Evidence on the Origin of the Haitian Outbreak*, American Society for Microbiology, 2(4):e00157-11, doi:10.1128/mBio.00157-11.

International Monetary Fund World Economic and Financial Surveys (2011). *World Economic Outlook (WEO). Tensions from the Two-Speed Recovery: Unemployment, Commodities and Capital Flows*.

Islam M., B. Drasar & D. Bradley (1990). *Long-term persistence of toxigenic Vibrio cholerae O1 in the mucilaginous sheath of a blue-green alga, Anabaena variabilis*. The Journal of tropical medicine and hygiene, 93, 133-139.

Jensen M. A., S. M. Faruque, J. J. Mekalanos & B. R. Levin (2006). *Modeling the role of bacteriophage in the control of cholera outbreaks. Proceedings of the National Academy of Sciences*, 103(12): 4652–4657.

King A. A., E. L. Ionides, M. Pascual & M. J. Bouma (2008). *Inapparent infections and cholera dynamics. Nature*, 454, doi:10.1038.

Kirn T. J., M. J. Lafferty, C. M. P Sandoe & R. K. Taylor (2000). *Delineation of pilin domains required for bacterial association into microcolonies and intestinal colonization. Molecular Microbiology*, Vol. 35(4):896-910.

Koelle K., X. Rodò, M. Pascual, Md. Yunus & G. Mostafa (2005). *Refractory periods and climate forcing in cholera dynamics. Nature*, 436, 696-700.

Lipp E. K., A. Huq & R. R. Colwell (2002). *Effects of global climate on infectious disease: The cholera model. Clinical Microbiology Reviews*, 15, 757–770.

Longini Jr. I. M., A. Nizam, M. Ali, M. Yunus, N. Shenvi & J. Clemens (2007). *Controlling endemic cholera with oral vaccines. PLoS Medicine*, 4, 1776-1783.

Mari L., E. Bertuzzo, L. Righetto, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe & A. Rinaldo (2011). *Modelling cholera epidemics: the role of waterways, human mobility and sanitation. Journal of The Royal Society Interface*. doi:10.1098/rsif.2011.0304.

MSPP (2010). *Report de cas – 23 October 2010. Ministère de la Santé Publique et de la Population. [On line] Disponible a:*

<<http://www.mspp.gouv.ht/site/downloads/Rapport%20jour>>.

MSPP (2012). *Report de cas – 22 Janvier 2012. Ministère de la Santé Publique et de la Population. [On line] Disponible a:*

<<http://www.mspp.gouv.ht/site/downloads/Rapport%20jour>>.

ONU (2010). [On line] Disponibile a:

<[http://ansa.it/web/notizie/rubriche/topnews/2010/02/24/visualizza\\_new.html\\_1707218513.html](http://ansa.it/web/notizie/rubriche/topnews/2010/02/24/visualizza_new.html_1707218513.html)>.

ORNL. Oak Ridge National Laboratory. [On line] Disponibile a:

<<http://www.ornl.gov/sci/landscan/index.shtml>>.

PAHO (2011). Pan American Health Organization. [On line] Disponibile a:

<[http://new.paho.org/hq/images/Atlas\\_IHR/CholeraHispaniola/atlas.html](http://new.paho.org/hq/images/Atlas_IHR/CholeraHispaniola/atlas.html)>.

PAHO (2012). Pan American Health Organization. *Experts call for major investments in water and sanitation to end cholera in Haiti and the Dominican Republic*. [On line] Disponibile a: <<http://new.paho.org/colera/?p=48>>.

Pascual M., K. Koelle & A. P. Dobson (2006). *Hyperinfectivity in cholera: a new mechanism for an old epidemiological model?*. PLOS Medicine, 3, 931-938.

Piarroux R., R. Barraï, B. Faucher, R. Haus, M. Piarroux, J. Gaudart, R. Magloire & D. Raoult (2011). *Understanding the Colera Epidemic, Haiti*. Emerging Infectious Diseases, Vol. 17, No. 7, 1161-1167.

Porporato A., G. Vico & P. A. Fay (2006). *Superstatistics of hydro-climatic fluctuation and interannual ecosystem productivity*. Geophysical Research Letters, 33, L15402, doi:10.1029/2006GL026412.

Righetto L., E. Bertuzzo, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe & A. Rinaldo (2010). *Modelling human movement in cholera spreading along fluvial systems*. Ecohydrology, 4 (1), 49-55, doi:10.1002/eco.122.

Righetto L., R. Casagrandi, E. Bertuzzo, L. Mari, M. Gatto, I. Rodriguez-Iturbe & A. Rinaldo (2012). *The role of aquatic reservoir fluctuations in a long-term cholera patterns*. Epidemics, 4, 33-42.

Rinaldo A., E. Bertuzzo, L. Mari, L. Righetto, M. Blokesch, M. Gatto, R. Casagrandi, M. Murray, S. Vesenbeckh & I. Rodriguez-Iturbe (2012). *Reassessment of the 2010-2011 Haiti cholera outbreak and rainfall-driven multi-season projection*. Proceedings of the National Academy of Sciences [in stampa].

Sack D. A., R. Bradley Sack, G. Balakrish Nair & A. Siddique (2004). *Cholera*. Lancet, 363, 223-233.

Tuite A., J. Tien, M. Eisenberg, D. Earns, J. Ma & D. Fisman (2011). *Cholera epidemic in Haiti, 2010: using a transmission model to explain spatial spread of disease and identify optimal control interventions*. Annals of Internal Medicine, 154, 593-601.

WHO (2008). *The World Health report 2008 – Primary health care (now more than ever)*. World Health Organization. Technical report.

WHO (2010a). *Prevention and control of cholera outbreaks: WHO policy and recommendations*. World Health Organization, Regional Office for the Eastern Mediterranean. Technical report.

WHO (2010b). *Cholera vaccines: WHO position paper*. World Health Organization. Weekly epidemiological record. 13, 2010, 85, 117-128. [On line] Disponibile a: <<http://who.int/wer>>.

WHO (2011). *Annual Report on Cholera – Cholera 2010*. World Health Organization. Weekly epidemiological record. 31, 2011, 86, 325-340. [On line] Disponibile a: <<http://www.who.int/wer/2011/wer8631.pdf>>.

WHO (2012). *Cholera, reporting outbreaks, 2010-2011*. World Health Organization. Global Health Observatory Map Gallery. [On line] Disponibile a: <[http://gamapserver.who.int/mapLibrary/Files/Maps/Global\\_Cholera\\_ITHRisk\\_2012\\_0118.png](http://gamapserver.who.int/mapLibrary/Files/Maps/Global_Cholera_ITHRisk_2012_0118.png)>.

Xia Y., O. N. Bjørnstad & B. T. Grenfell (2010). *Measles metapopulation dynamics, a gravity model for epidemiological coupling and dynamics*. *The American Naturalist*, 4 (164), 267-281.



## RIASSUNTO

Il colera colpisce ogni anno in media 178 000-327 000 persone, causando 4 000-6 300 decessi. Tuttavia, spesso i casi imputabili al colera non sono identificati e vengono classificati come diarrea acuta. Infatti la WHO stima che il numero reale dei casi possa aggirarsi intorno ai 3-5 milioni e quello dei decessi intorno a 100 000-130 000 all'anno [WHO, 2010a]. Dal 2009 al 2010 i casi registrati sono aumentati del 43% a livello globale e del 130% rispetto al 2000 [WHO, 2011].

La più recente epidemia è stata registrata ad Haiti nell'ottobre 2010 e ha avuto il suo massimo a gennaio 2011. A causa del terremoto di magnitudo 7.0  $M_w$  che ha colpito Haiti il 12 gennaio 2010 (il cui epicentro è stato localizzato a circa 25 km dalla capitale Port-au-Prince), la maggior parte delle infrastrutture, incluse quelle idriche, è andata distrutta. Questa situazione ha favorito la diffusione del colera, malattia collegata principalmente ad un difficile accesso all'acqua potabile e ad una carenza delle condizioni igienico-sanitarie. Il primo caso di colera riportato ad Haiti risale al 21 ottobre 2010, vicino a Mirebalais, nel distretto di Centre [Piarroux *et al.*, 2011]. A metà gennaio 2012, sono stati osservati 520 000 casi di colera, dei quali 7005 hanno portato al decesso [PAHO, 2012].

Il colera, che è causato dal batterio *Vibrio cholerae*, è una malattia che si trasmette per via oro-fecale. Una volta nell'intestino umano il batterio trova le condizioni ideali per potersi riprodurre, provocando diarrea e vomito nell'individuo ospite. Questi sintomi possono causare una severa disidratazione e, nei casi peggiori, la morte. Generalmente il colera può essere trattato con successo attraverso la reidratazione orale o, nei casi più gravi, per via endovenosa. Il tempo di incubazione varia dalle 2 ore ai 5 giorni, il che in parte spiega l'andamento esplosivo delle epidemie [WHO, 2010b]. In molti casi gli infetti non mostrano alcun sintomo (circa

l'80%). I batteri sono presenti nelle feci dei malati (sintomatici e asintomatici) e possono quindi essere espulsi nell'ambiente esterno, che viene così contaminato.

La veloce diffusione della malattia e il numero di persone contagiate mostrano l'importanza di sviluppare un metodo adeguato in grado di prevedere e ridurre l'impatto dell'epidemia. A questo scopo è necessario implementare un opportuno modello quantitativo che permetta lo sviluppo di politiche di pubblica sanità. Poiché è stata osservata una relazione tra l'aumento dei casi di infetti e le precipitazioni (infatti la risorgenza del colera è coincisa proprio con la stagione delle piogge), nella presente tesi viene validato un modello epidemiologico che considera le precipitazioni come una forzante esterna, che favorisce la diffusione del colera. Di conseguenza, per prevedere l'evoluzione dell'epidemia di colera che sta colpendo Haiti, è necessario eseguire delle previsioni di precipitazione, da inserire come ingresso nel modello usato per le simulazioni su breve e lungo periodo. Vengono proposti diversi metodi per generare la precipitazione e si confrontano i diversi scenari ottenuti.

## **IL MODELLO**

Il modello epidemiologico impiegato è basato sul modello spazialmente esplicito proposto da *Rinaldo et al.* (2012), che illustra le dinamiche tra quattro compartimenti (suscettibili, infetti, ristabiliti e batteri) e descrive la diffusione di *Vibrio cholerae* attraverso il trasporto idrologico e la mobilità umana, considerando anche l'effetto della precipitazione. Rispetto ai modelli precedenti, nel modello utilizzato in questa tesi si introducono le seguenti innovazioni [*Rinaldo et al.*, 2012]:

- si introduce la perdita di immunità, che è un flusso dal compartimento dei ristabiliti a quello dei suscettibili
- si include il ruolo del trasporto dovuto alla mobilità umana, considerando gli spostamenti sia degli individui suscettibili che di quelli infetti tra le diverse comunità

- la precipitazione è esplicitamente considerata come una forzante esterna che modifica il tasso di contaminazione dell'acqua dal batterio *Vibrio cholerae*.

### **Il modello epidemiologico spazialmente esplicito**

Quando si studiano situazioni a grande scala, come nel caso di Haiti, per meglio descrivere le dinamiche spazio-temporali è opportuno utilizzare un modello spazialmente esplicito. In questo tipo di modello il territorio viene diviso in unità fondamentali. Per descrivere le connessioni tra queste unità si utilizza un grafo i cui nodi sono costituiti dalle unità fondamentali e gli archi rappresentano le connessioni tra le diverse unità. Per ogni nodo si sviluppa un modello epidemiologico locale che viene accoppiato con il modello di trasporto di *Vibrio cholerae* tra le diverse comunità. Si suppone che i meccanismi principali di diffusione del batterio siano essenzialmente il trasporto idrologico e la mobilità umana.

Per il caso di studio sono state scelte come unità fondamentali in cui dividere il territorio delle entità idrologiche (bacini). Per l'estrazione dei bacini, che costituiscono i nodi della rete, si utilizzano mappe digitali di elevazione del terreno (DEM) dell'isola di Hispaniola, forniti da *United States Geological Survey* (USGS). Per questo lavoro è stata considerata l'area che impatta idrologicamente su Haiti e quindi, per preservare la continuità idrologica dei corsi d'acqua, è stato necessario includere anche parte delle Repubblica Dominicana. Il risultato di questo processo è la divisione del territorio in 366 bacini, ognuno dei quali costituisce un nodo della rete [Biermann, 2011]. A ogni nodo sono assegnate le coordinate del centro di massa della distribuzione della popolazione del corrispondente bacino.

#### *Il trasporto idrologico*

Il *Vibrio cholerae* può sopravvivere nell'ambiente acquatico anche in associazione con lo zooplancton o con la vegetazione acquatica [Colwell, 1996] e quindi si propaga lungo la rete idrologica anche insieme a questi organismi. Per

descrivere il trasporto idrologico, ogni nodo è connesso agli altri coerentemente con la rete idrologica.

### La mobilità umana

La mobilità umana è basata sui movimenti giornalieri della popolazione. I meccanismi di contaminazione associati al movimento umano sono due: da una parte gli individui suscettibili appartenenti a nodi non infetti possono contrarre il colera durante lo spostamento in altre comunità e, una volta tornati nel proprio nodo di origine, possono svolgere il ruolo di agenti contaminanti della propria riserva idrica; dall'altra, gli individui infetti (generalmente asintomatici) possono diffondere il batterio durante i loro spostamenti quotidiani in comunità differenti dalla propria. Infatti l'alto rapporto tra asintomatici e sintomatici rende ragionevole ipotizzare che il movimento delle persone affette da colera non sia sempre impedito.

A causa della conformazione di Haiti, in questo contesto non è realistico considerare la distanza euclidea tra due punti generici (per esempio per nodi separati dal mare). Per questa ragione, è stato prima creato un grafo nel quale ogni nodo è collegato soltanto con i quattro più vicini con distanze euclidee. La distanza tra due nodi generici  $i$  e  $j$  è calcolata come la lunghezza del cammino più breve, lungo il grafo precedente, che connette  $i$  e  $j$ . Il cammino più breve è stato trovato con l'algoritmo standard Dijkstra. La rete di mobilità umana è dunque definita attraverso un grafo completamente connesso.

### La precipitazione

Poiché la seconda recrudescenza della malattia avviene in corrispondenza dell'inizio della stagione delle piogge, è plausibile presumere che sia presente una relazione tra l'aumento dei casi di colera e la precipitazione (**Figura 1**). In seguito a questa considerazione si introduce nel modello una descrizione dell'aumento del tasso di contaminazione da *Vibrio cholerae* causato dall'incremento dell'intensità di precipitazione.

## Il modello finale

Il modello SIRB comprendente i termini di trasporto idrologico, di mobilità umana e la precipitazione, risulta avere la seguente formulazione:

$$\left\{ \begin{array}{l} \frac{dS_i}{dt} = \mu(H_i - S_i) - \beta \left[ (1-m) \frac{B_i}{K+B_i} + m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} \right] S_i + \rho R_i \\ \frac{dI_i}{dt} = \beta \left[ (1-m) \frac{B_i}{K+B_i} + m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} \right] S_i - (\gamma + \alpha + \mu) I_i \\ \frac{dR_i}{dt} = \gamma I_i - (\rho + \mu) R_i \\ \frac{dB_i}{dt} = -\mu_B B_i - l \left( B_i - \sum_{j=1}^n P_{ji} \frac{W_j}{W_i} B_j \right) + \frac{p}{W_i} [1 + \lambda J_i(t)] [(1-m) I_i + m \sum_{j=1}^n Q_{ji} I_j] \end{array} \right.$$

Il sistema è composto da  $4 \cdot n$  equazioni, dove  $n$  è il numero dei nodi.

L'evoluzione del compartimento dei suscettibili (prima equazione) è descritta dal bilancio tra la demografia della popolazione, gli individui contagiati e i guariti che perdono l'immunità. La popolazione umana è assunta, in assenza dell'epidemia, all'equilibrio demografico ( $H_i$ ). Il parametro  $\beta$  rappresenta il tasso di esposizione all'acqua contaminata, mentre  $B_i/(K+B_i)$  è una funzione saturante che indica la probabilità di diventare infetti a causa di un'esposizione a una concentrazione di batteri  $B_i$ . In questo formalismo,  $K$  è la costante di semisaturazione, cioè la concentrazione di patogeni per la quale la probabilità di un suscettibile di diventare infetto è la metà del valore massimo. Il parametro  $m$  indica la parte di popolazione di un nodo che viaggia quotidianamente e raggiunge il nodo di destinazione  $j$  con una probabilità  $Q_{ij}$  e poi torna al proprio nodo di partenza. Viceversa  $1-m$  è la parte di popolazione che rimane stanziale nel nodo. In analogia col modello gravitazionale, la probabilità  $Q_{ij}$  è definita come:

$$Q_{ij} = \frac{H_j e^{-\frac{d_{ij}}{D}}}{\sum_{k \neq i}^n H_k e^{-\frac{d_{ik}}{D}}}$$

dove  $H_j$  e  $H_k$  sono le popolazioni rispettivamente del nodo  $j$  e del nodo  $k$ ,  $d_{ij}$  e  $d_{ik}$  sono le distanze tra i nodi  $i$  e  $j$  e tra i nodi  $i$  e  $k$ . Infine  $D$  è la distanza di dispersione

media. In questo modello il nodo  $j$  è tanto più attrattivo per il nodo  $i$  quanto più è grande la popolazione di  $j$  e tanto più è piccola la distanza che li separa ( $d_{ij}$ ). La capacità di movimento è differente da nazione a nazione e dipende dalle infrastrutture, dalla topografia e dalle abitudini locali, quindi il parametro  $D$  deve essere calibrato a seconda delle situazioni. Il termine  $\beta m \sum_{j=1}^n Q_{ij} \frac{B_j}{K+B_j} S_i$  è la frazione di suscettibili che contrae la malattia viaggiando.

La dinamica del compartimento degli infetti (seconda equazione) è data dal bilancio tra il flusso entrante di nuovi malati e quello uscente dovuto alle guarigioni e alla mortalità naturale, considerando anche un tasso di mortalità aggiuntivo dovuto al colera ( $\alpha$ ).

La terza equazione, che descrive la dinamica dei ristabiliti, è data dal bilancio tra i guariti dal colera e coloro che muoiono naturalmente o che perdono l'immunità. Il parametro  $\rho$  quantifica questa perdita dell'immunità e, quindi, la relativa ripopolazione del serbatoio dei suscettibili.

L'ultima equazione descrive la concentrazione di batteri che vivono nella riserva d'acqua locale. I batteri che vivono nell'ambiente naturale hanno un tasso di mortalità  $\mu_B$ . Il termine  $l$  è il tasso di mobilità dei batteri (espresso in giorni<sup>-1</sup>) che si muovono per il trasporto idrologico dal nodo  $i$  al nodo ad esso interconnesso  $j$  con una probabilità pari a  $P_{ij}$ . In particolare,  $1/l$  è il tempo di residenza medio dei batteri in ciascuna riserva d'acqua. Nel caso di studio si assume che il trasporto avvenga solo da monte a valle. Il termine  $\frac{p}{W_i} m \sum_{j=1}^n Q_{ji} l_j$ , rappresenta la quantità di batteri introdotta nel nodo  $i$  da tutti gli infetti asintomatici appartenenti ad altri nodi, che viaggiano quotidianamente verso il nodo  $i$  [Bertuzzo *et al.*, 2011]. Infatti, gli infetti contribuiscono ad aumentare la concentrazione di vibrioni con un tasso  $p/W_i$ , dove  $p$  è il tasso di batteri escreti da un individuo infetto che raggiunge e contamina la riserva d'acqua locale di volume  $W_i$ . Esso è considerato essere proporzionale alla popolazione  $W_i=cH_i$ , in quanto si presuppone che una città con più abitanti abbia a disposizione riserve d'acqua maggiori [Mari *et al.*, 2011]. Per giustificare la stagionale risorgenza dell'epidemia di colera, si assume che il tasso di

contaminazione dell'acqua  $p$  possa aumentare con la precipitazione  $J_i(t)$ , attraverso un coefficiente  $\lambda$ . Questo è dovuto al fatto che la precipitazione causa un aumento dell'agente patogeno nella riserva d'acqua, in quanto attraverso il deflusso superficiale della precipitazione vengono mobilitati grandi quantità di batteri dai siti di defecazione all'aria aperta verso corpi idrici utilizzati come riserva potabile e per uso domestico [Rinaldo et al., 2012].

Le condizioni iniziali dei nodi infetti sono  $X_i(0) = [S_{i0}, I_{i0}, 0, B_{i0}]^T$ , con  $S_{i0} = H_i - I_{i0}$ , dove  $I_{i0}$  è il numero di infetti iniziali, osservati nella realtà, e  $B_{i0}$  è la concentrazione di batteri iniziale, calcolata come la concentrazione di quasi equilibrio  $B_{i0} = \frac{pI_{i0}}{W_i\mu_B}$ . Invece, le condizioni iniziali dei nodi sani sono  $X_i(0) = [H_i, 0, 0, 0]^T$ , poiché non sono presenti individui infetti. Il modello presuppone la conoscenza della popolazione presente in ogni nodo. I dati relativi alla popolazione sono in formato raster e sono stati forniti dall'*Oak Ridge National Laboratory* (ORNL), che distribuisce i dati del censimento, raccolti secondo i confini amministrativi, sul territorio, considerando la copertura del terreno dell'area di interesse. In particolare, nel momento in cui si suddivide la popolazione, ad ogni copertura viene attribuito un peso diverso. I dati hanno approssimativamente  $1 \text{ km}^2$  di risoluzione ( $30 \times 30$  arcosecondi, ovvero  $0.0083 \times 0.0083$  gradi) e rappresentano la distribuzione della popolazione mediata su 24 ore (ovvero rilevando i dati di distribuzione della popolazione per ogni ora della giornata ed effettuandone la media) [ORNL].

Ogni modello locale presenta, sotto certe condizioni, oltre all'equilibrio corrispondente alla popolazione sana alla capacità portante  $\bar{X}_H = [H, 0, 0, 0]^T$ , un secondo equilibrio corrispondente alla popolazione parzialmente infetta  $\bar{X}_+ = [\bar{S}_+, \bar{I}_+, \bar{R}_+, \bar{B}_+]^T$ . Quest'ultimo equilibrio ha però senso biologico solo se  $\bar{I}_+$  non è negativo, quindi deve essere verificata la condizione  $R_0 \geq 1$ , dove  $R_0$  è il tasso base di riproduzione della malattia, che può essere interpretato come il numero medio di infezioni secondarie prodotte da un individuo infetto immesso in una popolazione sana e costituita da  $H$  individui ed è definito come:

$$R_0 = \frac{\beta p H}{(\gamma + \alpha + \mu) K \mu_B W}$$

Se  $R_0$  è minore di 1, l'equilibrio con la popolazione parzialmente infetta non può sussistere. Si può dimostrare che in questo caso l'equilibrio  $\bar{X}_H$  è stabile qualsiasi siano le condizioni iniziali e quindi l'epidemia va estinguendosi naturalmente. Invece, se  $R_0$  è maggiore di 1, l'equilibrio corrispondente alla popolazione parzialmente infetta  $\bar{X}_+$  è l'unico equilibrio stabile e quindi la malattia è endemica, ovvero una frazione costante della popolazione è infetta [Codeço, 2001].

Il modello descritto simula l'andamento dell'epidemia, considerando come infetti sia gli individui sintomatici che quelli asintomatici, ma in realtà i dati disponibili riguardano soltanto i casi osservati (che sono quelli sintomatici). Per poter confrontare le simulazioni effettuate con i dati reali, si definisce il numero di casi cumulati simulati sintomatici per ogni nodo:

$$\frac{dC_i}{dt} = \sigma \cdot \beta \left[ (1 - m) \frac{B_i}{K + B_i} + m \sum_{j=1}^n Q_{ij} \frac{B_j}{K + B_j} \right] S_i$$

Il fattore  $\sigma$  corrisponde alla frazione di individui infetti che, presentando i sintomi del colera, si rivolgono a strutture ospedaliere (circa il 20% degli infetti). Il secondo fattore rappresenta il flusso di nuovi individui infetti, calcolati sommando tutte le persone che non viaggiano e che quindi sono state contagiate a causa dei batteri presenti nella propria comunità, e tutte quelle che sono state infettate durante il loro viaggio giornaliero (e quindi da batteri presenti in altri nodi).

### Regime pluviometrico

Le statistiche di precipitazione sono calcolate sulla base dei dati forniti da:

- *NASA-JAXA's Tropical Rainfall Measuring Mission* attraverso il geoportale FEWS NET, messo a disposizione da USGS, che fornisce dati decadali, dal 1 giugno 2010 fino al 30 settembre 2011.



- dal geoportale <http://portal.iri.columbia.edu/portal/server.pt> dell'*International Research Institute for Climate and Society*, è possibile reperire dati decadal di precipitazione, dal 1 gennaio 2002 al 31 luglio 2011 (**Figura 2**), e dati di precipitazione giornalieri che si riferiscono al periodo dal 1 gennaio 2002 fino al 31 ottobre 2011.

Per quanto riguarda i dati decadal, ogni mese è suddiviso in tre decadi, di cui l'ultima di lunghezza variabile. In entrambi i casi, i dati utilizzati sono spazialmente distribuiti, con una risoluzione di 0.25 gradi di latitudine e longitudine. Ogni immagine è composta da 12x27 celle, cosicché l'area totale coperta è pari a circa 330x750 km<sup>2</sup>.

Per trasferire le precipitazioni dalle celle ai nodi o ai dipartimenti, si assegna ad ogni bacino la precipitazione calcolata con l'algoritmo *nearest neighbor* eseguito sulle precipitazioni delle celle che coprono il bacino in considerazione. In questo modo si ottengono dei valori puntuali di precipitazione per le entità desiderate.

Il regime pluviometrico di Haiti presenta una significativa variabilità inter-annuale e intra-annuale, con un alternarsi di periodi particolarmente piovosi ad altri particolarmente secchi. Il clima di Haiti è caratterizzato da due distinte stagioni piovose e da una stagione secca, tra dicembre e marzo (**Figura 3**). Il periodo tra giugno e novembre corrisponde alla stagione degli uragani, che si abbattano sull'isola con maggiore probabilità tra la fine di agosto e settembre.

### **Stima e calibrazione dei parametri**

Alcuni dei parametri utilizzati nel modello sono stati ricavati dalla letteratura esistente, mentre altri sono stati stimati tramite calibrazione effettuata dal *Laboratory of Ecohydrology* dell'*Ecole Polytechnique Fédérale de Lausanne* (**Tabella 1**). Per questi ultimi, si utilizza un approccio che si basa sul campionamento *Markov Chain Monte Carlo* (MCMC).

La calibrazione è stata eseguita per il periodo dal 23 ottobre 2010 al 28 maggio 2011. Per i dati relativi ai casi di colera sono stati utilizzati quelli forniti dal *Ministère de la Santé Publique et de la Population* (MSPP) di Haiti e dalla *Pan*

*American Health Organization (PAHO)*, mentre per i dati di precipitazione sono stati utilizzati i dati decadali forniti dalla *NASA-JAXA's Tropical Rainfall Measuring Mission*.

Parametro	Valore	Unità di misura	Tipo
$\mu$	$1/(61.4 \cdot 365)$	giorni <sup>-1</sup>	Da letteratura
$\beta$	1.0	giorni <sup>-1</sup>	Da letteratura
$\rho$	$1/(3 \cdot 365)$	giorni <sup>-1</sup>	Da letteratura
$\gamma$	0.20	giorni <sup>-1</sup>	Da letteratura
$\alpha$	$4.0 \cdot 10^{-3}$	giorni <sup>-1</sup>	Da letteratura
$\mu_B$	0.2	giorni <sup>-1</sup>	Da letteratura
$\sigma$	0.2	-	Da letteratura
$R_0$	4.9174	-	Calibrato
$L$	6.5501	giorni <sup>-1</sup>	Calibrato
$M$	0.6761	-	Calibrato
$\lambda$	2.5154	giorni /mm	Calibrato
$D$	100	km	Calibrato
$\theta$	0.2007	giorni <sup>-1</sup>	Calibrato

**Tabella 1:** Riepilogo dei parametri utilizzati.

## VALIDAZIONE DEL MODELLO EPIDEMIOLOGICO E PREVISIONE DELL'EVOLUZIONE DEL COLERA

Per effettuare le simulazioni dell'evoluzione dell'epidemia di colera ad Haiti, si devono prima generare le previsioni di precipitazione. Infatti queste ultime costituiscono uno degli ingressi del modello SIRB utilizzato nella presente tesi.

### Validazione con precipitazioni reali

Avendo a disposizione i dati reali di precipitazione fino al 30 settembre 2011, con questi si realizza una prima simulazione del modello epidemiologico che può essere considerata come una validazione.

Innanzitutto è possibile osservare dalla **Figura 4** che per il periodo relativo alla calibrazione, l'andamento dei casi simulati (linea rossa) rispecchia quello dei casi registrati (linea nera). Per il periodo di validazione, si riesce a descrivere l'aumento di casi settimanali di colera nel periodo tra fine maggio e giugno 2011, ma il modello non riesce poi a riprodurre bene il rapido declino dei casi dopo il picco di metà giugno. Si noti anche come i casi di colera siano sovrastimati. In generale però è preferibile che il modello sovrastimi i casi piuttosto che il contrario, in modo da non sottovalutare la portata dell'evoluzione dell'epidemia.

Nella simulazione, escluso il periodo di forte diffusione iniziale, un aumento dei casi di colera è generalmente anticipato da un periodo di intensa precipitazione. Per esempio, per la prima decade di giugno si registra un picco di precipitazione e nelle due settimane seguenti si osserva un aumento dei casi registrati. Infatti, bisogna considerare che l'effetto dovuto alle precipitazioni sull'aumento del contagio non è istantaneo, in quanto il tempo di incubazione della malattia può durare qualche giorno. Di conseguenza le simulazioni risultano essere attendibili e non distanti dalla realtà, ad indicare un'ottima prestazione del modello di previsione dell'epidemia di colera.

### **Simulazione con l'anno di precipitazione medio**

Come primo metodo per proiettare in avanti le precipitazioni a partire dal 29 maggio 2011, si propone l'utilizzo delle precipitazioni medie, calcolate sulla base dei dieci anni di dati decadal (dal 2002 al 2011) forniti dall'*International Research Institute for Climate and Society* e opportunamente risolte alla scala dei nodi. Si esegue una simulazione su lungo periodo a partire dal 23 ottobre 2010, utilizzando i dati reali di precipitazione fino al 28 maggio 2011 e quelli dell'anno medio dal 29 maggio 2011 al 31 dicembre 2013, in modo da poter prevedere eventuali risorgenze dell'epidemia. Oltre all'anno medio, si utilizzano anche il 25-esimo percentile delle precipitazioni di ogni nodo (quindi si impone che su tutta Haiti la precipitazione sia particolarmente scarsa) e il 75-esimo percentile di ogni nodo (quindi si impone che la

precipitazione sia abbondante dappertutto). In **Figura 5** sono rappresentati i risultati ottenuti con questo metodo.

Per quanto riguarda il periodo dal 29 maggio al 30 settembre 2011 è possibile confrontare direttamente i risultati della simulazione con i dati. Anche nella simulazione effettuata utilizzando la precipitazione dell'anno medio, si riesce a prevedere l'incremento dei casi di colera nel periodo maggio-giugno. Nei tre mesi successivi i casi simulati con l'anno medio di precipitazione sono però fortemente sovrastimati rispetto a quelli reali. Questa differenza è causata dalla maggiore intensità delle precipitazioni medie rispetto a quelle reali. Infatti, sebbene l'anno di precipitazione medio preveda un aumento di precipitazione verso agosto-settembre, il 2011 è risultato essere un anno particolarmente secco, in cui non è presente la seconda stagione piovosa tipica di Haiti. Si osservi come nei mesi di agosto e di settembre 2011 la precipitazione relativa sia inferiore anche rispetto al primo quartile delle precipitazioni (**Figura 5c**). Questo fatto può rappresentare una possibile spiegazione della diminuzione dei casi di colera registrati rispetto a quelli simulati, che prevedono in corrispondenza del picco di precipitazione di fine agosto, un ulteriore aumento dei casi di colera settimanali (**Figura 5a**).

Venendo al periodo di previsione futura, è possibile mettere in evidenza le seguenti caratteristiche. In corrispondenza con la stagione secca (da novembre 2011 a aprile 2012), l'epidemia sembra estinguersi, con un centinaio di casi a settimana in tutta Haiti. Si prevede una risorgenza dei casi di colera nel periodo tra agosto e novembre 2012, dopo le due stagioni delle piogge, con un picco di circa 8700 casi settimanali previsto per il 13 ottobre 2012. Il contagio risulta comunque essere contenuto rispetto al 2011. Secondo la simulazione effettuata, per l'orizzonte di previsione, si verifica anche una terza risorgenza della malattia nel periodo tra luglio e novembre 2013, che coinvolge più persone rispetto a quella del 2012. Infatti, stabilendo che la perdita di immunità avviene in media dopo 3 anni, chi aveva contratto il colera nel 2010, da settembre 2013 ritorna mediamente a essere nuovamente suscettibile, ripopolando così il bacino delle persone potenzialmente infettabili.

### **Simulazione con la precipitazione dell'anno tipico**

Invece della precipitazione dell'anno medio, per il periodo dal 29 maggio 2011 al 31 dicembre 2013 è possibile utilizzare la precipitazione relativa all'anno tipico come ingresso nel modello di previsione dell'epidemia di colera. L'anno tipico è l'anno osservato i cui valori meno si discostano dall'anno medio. Studiando i dati decadalni forniti dall'*International Research Institute for Climate and Society* per gli anni dal 2002 al 2010, è possibile individuare il 2002 come anno tipico, in quanto la differenza tra la precipitazione del 2002 e quella dell'anno medio presenta lo scarto minore. I risultati ottenuti sono rappresentati in **Figura 6**.

L'andamento delle precipitazioni dell'anno tipico presenta dei picchi più alti e dei minimi più accentuati rispetto a quelli dell'anno medio (**Figura 6c**), causando una risposta notevolmente diversa della malattia. Infatti si nota subito un picco di casi di colera a inizio settembre 2011 (**Figura 6a**). Entrando più nel dettaglio, l'anno tipico segue meglio l'andamento delle precipitazioni registrate nella realtà rispetto all'anno medio per le prime due decadi del mese di giugno, ma per quanto riguarda la terza decade di giugno, la precipitazione giornaliera dell'anno tipico è di 1.37 mm/giorno, mentre per il 2011 risulta essere di 4.55 mm/giorno. Quindi il fatto che l'andamento dei casi di colera simulati con l'anno tipico segua quello dei casi reali non deve ingannare, dato che si tratta di due ingressi di precipitazioni differenti. Da luglio 2011 la precipitazione dell'anno tipico presenta un forte incremento e risulta essere sempre superiore non solo rispetto alla precipitazione reale, ma anche rispetto a quella media. Questo provoca una rapida diffusione dell'epidemia, con un picco di casi settimanali pari a 39 449 il 3 settembre 2011. Si noti come questo massimo si verifichi poco dopo il picco di precipitazione della terza decade di agosto, ad indicare come il modello risponda sensibilmente alle variazioni di precipitazione. Un aumento così repentino dei casi di colera potrebbe essere spiegato dal fatto che maggiori fluttuazioni nelle precipitazioni causano una risposta più impulsiva nella diffusione della malattia.

### **Simulazione con la precipitazione generata con un generatore puntuale di precipitazione poissoniano**

Un raffinamento ulteriore per la generazione degli afflussi meteorici è quello di usare un generatore puntuale (alla scala dei distretti o dei bacini) di precipitazione basato su un processo di tipo poissoniano (*point rainfall Poisson's generator*). La generazione è di tipo puntuale e quindi non considera in modo esplicito le correlazioni spaziali.

In questo generatore si suppone che sia l'intensità della precipitazione sia il tempo di inter-arrivo abbiano una distribuzione esponenziale:

$$f(x) = \gamma e^{-\gamma x} \quad f(\tau) = \lambda e^{-\lambda \tau}$$

dove  $x$  è l'intensità della precipitazione,  $1/\gamma$  il valore medio dell'intensità della precipitazione e  $\tau$  è il tempo di inter-arrivo con valor medio  $1/\lambda$ .

Sebbene siano disponibili i dati di precipitazione anche relativi a parte del 2011, è possibile utilizzare soltanto gli anni completi, e quindi si lavora con soli nove anni di dati (2002-2010). Innanzitutto si devono calcolare le grandezze su cui si basa la generazione, che sono:

- Tasso di arrivo dei giorni piovosi in un mese ( $\lambda$ )
- Precipitazione media mensile (contando solo i giorni piovosi)

La procedura per stimare i parametri dei generatori è quella di far coincidere i valori teorici dei parametri dei generatori con quelli campionari. Utilizzando i generatori si originano cento anni di precipitazione per ogni nodo o dipartimento (in **Figura 7** si riporta a titolo d'esempio un confronto tra le statistiche relative alla precipitazione generata rispetto a quelle della precipitazione reale del distretto Nord-Ouest). L'obiettivo è quello di inserire la precipitazione come ingresso nel modello di colera, e quindi, dato che si vogliono effettuare simulazioni a lungo termine (fino al 31 dicembre 2013), è necessario avere una sequenza di tre anni di dati di precipitazione. Per creare questa sequenza, ogni anno originato con il generatore è seguito da un anno estratto casualmente dai cento generati e così anche per il terzo (dove l'estrazione avviene sempre tra ai cento anni generati). In questo modo si

creano cento sequenze casuali di tre anni dagli anni generati dal generatore puntuale di tipo poissoniano. È possibile utilizzare il generatore sia a partire dalla precipitazione di ogni distretto che da quella di ogni nodo.

### **Simulazione con la precipitazione generata con un generatore puntuale di precipitazione poissoniano a partire dai distretti**

#### Generazione per dipartimenti disaggregati

Inizialmente si originano i 100 anni di precipitazioni basandosi sui nove anni di precipitazioni giornaliere dei distretti fornite dall'*International Research Institute for Climate and Society*. Dopodiché, poiché il modello epidemiologico lavora sui nodi, si effettua il *downscaling* della pioggia dai distretti ai bacini, considerando la frazione di ognuno di essi presente in ogni distretto. I risultati sono riportati in **Figura 8**.

Le simulazioni eseguite con la precipitazione originata dal generatore puntuale di precipitazione di Poisson seguono quella effettuata con la precipitazione media. Questo fatto potrebbe essere spiegato ricordando che, come mostrato in **Figura 7**, la media mensile degli anni generati e quella calcolata sui dati reali a disposizione corrispondono, così come il numero di eventi e la media delle precipitazioni nei giorni piovosi.

#### Generazione per dipartimenti aggregati

In precedenza i distretti sono stati considerati indipendenti tra loro. Tuttavia diverse aree di Haiti hanno un regime pluviometrico simile. Pertanto, si introducono delle aggregazioni tra distretti con caratteristiche simili nei *pattern* di precipitazione, definendo così un generatore di precipitazione per ogni gruppo di dipartimenti e non più per ogni singolo distretto. In riferimento agli andamenti medi mensili della precipitazione e alle osservazioni relative alla correlazione spaziale della precipitazione (**Figura 9**), il territorio di Haiti è stato suddiviso in cinque zone (**Figura 10**):

- Nord-Ouest
- Nord e Nord-Est (**Figura 11a**)
- Artibonite e Ouest (**Figura 11b**): presentano medie paragonabili e hanno lo stesso andamento delle precipitazioni medie mensili con un picco di precipitazioni in agosto
- Centre
- Sud-Est, Sud, Nippes e Grande Anse: presentano un regime pluviometrico paragonabile (**Figura 12**), con l'evidenza di due stagioni piovose, una a maggio e l'altra in agosto.

Si calcola per ogni gruppo la media della precipitazione giornaliera per ogni giorno dall'1 gennaio 2002 al 31 dicembre 2010 sulla base dei dati giornalieri forniti dall'*International Research Institute for Climate and Society*. Si ottengono quindi cinque serie che vengono inserite come ingresso nel generatore puntuale di precipitazioni di Poisson, che calcola cento anni di precipitazione per ogni gruppo. Successivamente, ad ogni distretto, si assegnano le precipitazioni generate del gruppo corrispondente. Infine, poiché i cento anni di precipitazione generati si riferiscono ai distretti, si scala la precipitazione dai distretti ai nodi, sui quali lavora il modello. I risultati sono rappresentati in **Figura 13**.

Rispetto al caso precedente, in cui la precipitazione era generata per ogni distretto, non si prevede il picco di casi settimanali a luglio 2011 (**Figura 13**). Il motivo per cui non sono presenti sostanziali differenze neanche a scala di distretto (**Figura 14** e **Figura 15**) può essere imputato al fatto che il generatore puntuale di precipitazione di Poisson si basa su serie storiche, che contengono già come informazione implicita la correlazione spaziale.

### **Simulazione con la precipitazione generata con un generatore puntuale di precipitazione poissoniano a partire dai bacini**

Si propone una generazione della precipitazione è che si basa sui nodi, in quanto sono le entità su cui lavora il modello di diffusione del colera. In questo caso si preserva, nel *downscaling* dei dati di partenza, la precipitazione specifica nel



territorio afferente al nodo, cosa non garantita quando si utilizza un generatore a livello di dipartimenti. Il procedimento è sempre lo stesso del caso precedente: dai dati di precipitazione registrati dal 2002 al 2010, si generano cento anni di piogge basandosi sulle precipitazioni dei nodi. Dopodiché si creano cento sequenze casuali di tre anni estratte dagli anni generati, in modo da avere una previsione dal 2011 al 2013. Si confrontano le simulazioni ottenute immettendo come ingresso nel modello epidemiologico la precipitazione giornaliera e quella decadale.

#### *Generazione di precipitazione partendo da dati giornalieri*

Per generare i cento anni di precipitazione si utilizzano le precipitazioni giornaliere fornite dall'*International Research Institute for Climate and Society*, per il periodo dal 1 gennaio 2002 al 31 dicembre 2010.

Dalla **Figura 16c** è possibile osservare che l'andamento delle precipitazioni generate (mediate sui nodi per ognuno dei cento anni generati) ha un andamento più costante rispetto a quello dell'anno di precipitazione medio. Per esempio, riferendosi al mese di maggio 2012, l'andamento dell'anno di precipitazione medio presenta una prima decade abbastanza secca, una seconda in linea con la media mensile e una terza molto piovosa. Questo comportamento è spiegabile ricordando che il generatore lavora su base mensile e quindi non considera la variabilità intra-mensile che, invece, è confermata essere presente ad Haiti dall'andamento dell'anno di precipitazione medio. Sarebbe quindi necessario utilizzare un generatore che adotti una scala temporale più fine di quella mensile. Inoltre, per generare insiemi di scenari più vari possibile, si potrebbe anche introdurre una variabilità inter-annuale, che vada a modificare il regime pluviometrico a livello stagionale, in quanto il modello epidemiologico si dimostra molto sensibile a queste variazioni. Infatti, la stagione particolarmente secca dell'estate 2011 è probabilmente responsabile del più rapido declino dei casi osservati rispetto alle simulazioni effettuate. In **Figura 16a** gli scenari generati con la precipitazione originata dal generatore non si discostano molto da quelli dell'anno di precipitazione medio. Questo fatto si spiega considerando che la generazione della precipitazione avviene indipendentemente per

ogni nodo, quindi per simulare situazioni di siccità in tutta l'isola o di estrema piovosità sarebbe necessario generare molti più anni di precipitazione rispetto ai cento attualmente generati, ma ne conseguirebbe un carico computazionale molto oneroso. Una possibile soluzione potrebbe essere quella di introdurre dei raggruppamenti dei bacini, in modo da imporre un regime precipitativo simile per alcune aree.

#### Generazione di precipitazione partendo da dati decadali

Si immettono come ingresso nel generatore puntuale di precipitazione di Poisson le precipitazioni decadali fornite dall'*International Research Institute for Climate and Society* per il periodo dal 1 gennaio 2002 al 31 dicembre 2010.

Osservando la **Figura 17c**, che riporta le medie sui nodi delle precipitazioni giornaliere, è possibile notare che la precipitazione generata devia in modo poco sensibile dalla media mensile e non presenta variazioni intra-mensili, che per essere colte richiederebbero l'adozione di una scala temporale più fine. Il limitato spettro di scenari prodotto (**Figura 17a**) è dovuto al fatto che, così come nel caso precedente, la generazione avviene indipendentemente per ogni nodo.

#### **Simulazione con la precipitazione calcolata con modelli ARX**

Tutti i metodi precedenti si basano su statistiche della serie storica di precipitazione. In questo paragrafo si propone invece un modello di previsione delle precipitazioni che, partendo dall'analisi della serie storica di precipitazione a disposizione, genera dei dati di pioggia considerando una consequenzialità temporale con i dati reali. Inoltre, si vorrebbe considerare anche la correlazione spaziale, poiché la precipitazione di zone confinanti è spesso correlata (**Figura 9**). È quindi necessario introdurre un modello di previsione di precipitazione che consideri sia la correlazione temporale che quella spaziale.

Per il caso di studio si propone l'utilizzo di un modello ARX (*Auto Regressive Exogenous*), in cui la correlazione temporale è descritta dalla parte auto-regressiva del modello, mentre quella spaziale dall'ingresso esogeno. La parte auto-regressiva

corrisponde alla precipitazione di un dato dipartimento registrata nei giorni precedenti, mentre l'ingresso esogeno corrisponde alle precipitazioni degli altri distretti. Per scegliere la complessità del modello ARX, ovvero per stabilire l'ordine per la regressione sia della precipitazione del distretto che per quella degli ingressi esogeni, sono stati calcolati gli AIC (*Akaike Information Criterion*) relativi a modelli con differenti combinazioni degli ordini.

Con questo metodo è possibile compiere delle previsioni di precipitazione per un dato distretto, utilizzando la precipitazione del distretto stesso e quella relativa agli altri distretti ai tempi precedenti. Attraverso un modello ARX è però possibile predire la precipitazione solo per la prima decade di simulazione (nel caso di Haiti dall'1 al 10 giugno 2011), poiché si dimostra che la precipitazione prevista a più di un passo di previsione risulta essere del tutto scorrelata dall'ingresso del modello.

Si esegue la previsione per i tre modelli migliori secondo il criterio AIC. Per la calibrazione dei modelli ARX sono state utilizzate le precipitazioni decadali fornite dall'*International Research Institute for Climate and Society*, dal 1 gennaio 2002 al 31 maggio 2011. Com'è possibile notare dalla **Figura 18** i tre modelli restituiscono una previsione di precipitazione molto simile. Dopo aver previsto una decade di precipitazione per tutti i distretti, si effettua una previsione dei casi di colera per la prima settimana di giugno. Come ingresso di precipitazione nel modello epidemiologico si utilizzano per il periodo dal 23 ottobre 2010 al 31 maggio 2011 le precipitazioni decadali della *NASA-JAXA's Tropical Rainfall Measuring Mission* e dal 1 giugno 2011 al 10 giugno 2011 la precipitazione prevista con il modello ARX sulla base delle precipitazioni decadali fornite dall'*International Research Institute for Climate and Society*. Con questo metodo di previsione della precipitazione, i casi di colera settimanali previsti per la prima settimana di giugno sono 9 264, considerando che nella realtà sono stati registrati 11 450 casi (**Figura 19**).

La previsione di precipitazione con il modello ARX comporta un miglioramento nella previsione a breve termine, in quanto segue l'andamento pluviometrico del periodo in corso, ma non permette di effettuare previsioni per orizzonti temporali maggiori di una decade. Comunque, il modello ARX risulta essere attendibile a

breve termine e può quindi essere impiegato in una politica di controllo in linea dell'epidemia, in cui le previsioni a dieci giorni permettono di implementare degli interventi su brevissimo periodo.

### **Confronto tra gli *output* epidemiologici ottenuti con i diversi metodi di previsione della precipitazione**

In generale è possibile affermare che il modello è molto sensibile all'ingresso di precipitazione, in quanto il numero di casi di colera aumenta con l'aumentare dell'intensità e delle fluttuazioni della precipitazione. Infatti, come per il caso dell'anno tipico, oscillazioni maggiori corrispondono a risposte più impulsive nella diffusione della malattia. Per quanto riguarda la previsione a lungo termine, tutte le simulazioni effettuate concordano nel prevedere una prima risorgenza dell'epidemia intorno a settembre 2012 e una seconda più acuta a settembre 2013.

In **Tabella 2** si riporta uno schema riassuntivo dei casi cumulati previsti per il 30 settembre 2011. Per ogni metodo sono anche mostrati la differenza con i casi reali e l'errore di stima.

Nonostante sia il metodo più semplice, tra i diversi metodi di previsione della precipitazione analizzati, quello che restituisce una maggiore variabilità degli scenari è il metodo che si basa sulle statistiche di precipitazione (media, 25-esimo percentile e 75-esimo percentile). Infatti, in questo caso si riescono a cogliere le situazioni estreme, in quanto si impone che su tutta Haiti la precipitazione sia particolarmente scarsa (25-esimo percentile) oppure particolarmente abbondante (75-esimo percentile). In particolare, la simulazione che meglio riesce a riprodurre il reale andamento dell'epidemia nel terzo trimestre del 2011 è quella effettuata con il 25-esimo percentile della precipitazione. Il motivo per il quale l'epidemia è meglio descritta dal 25-esimo percentile della precipitazione, piuttosto che dalla simulazione eseguita con l'anno di precipitazione medio, è che il 2011 risulta essere un anno particolarmente secco (quindi più vicino al 25-esimo percentile che alla media), in cui la stagione piovosa si è verificata solo nel periodo limitato tra maggio e luglio, quando in genere si protrae per un intervallo più lungo (da maggio fino a ottobre).

				Casi cumulati previsti (casi)	Differenza rispetto ai casi reali (casi)	Errore 1 <sup>11</sup> (%)	Errore 2 <sup>12</sup>
<b>CASI CUMULATI REALI REGISTRATI (al 30 settembre 2011)</b>				452 710	-	-	-
<b>Simulazione con precipitazione reale</b>				518 000	65 290	14	2.86E+10
<b>Simulazione con statistiche di precipitazione</b>	media			647 220	194 510	43	5.14E+10
	25-esimo percentile			485 000	32 290	7	2.84E+10
	75-esimo percentile			772 000	319 290	71	1.12E+11
<b>Simulazione con precipitazione dell'anno tipico</b>				667 440	214 730	47	5.09E+10
<b>Simulazione con il generatore puntuale di precipitazione di Poisson</b>	distretti	distretti disaggregati	caso MIGLIORE	582 980	130 270	29	3.47E+10
			caso PEGGIORE	692 310	239 600	53	7.66E+10
		distretti aggregati	caso MIGLIORE	572 990	120 280	27	3.39E+10
			caso PEGGIORE	684 810	232 100	51	6.78E+10
	nodi	precipitazioni giornaliere	caso MIGLIORE	624 060	171 350	38	4.50E+10
			caso PEGGIORE	657 030	204 320	45	5.81E+10
		precipitazioni decadali	caso MIGLIORE	633 560	180 850	40	4.82E+10
			caso PEGGIORE	660 240	207 530	46	5.83E+10

**Tabella 2:** Per ogni metodo di previsione di precipitazione è stato riportato il numero di casi di colera cumulati previsti dal modello epidemiologico per il 30 settembre 2011, la differenza rispetto a quelli reali e l'errore di stima commesso, calcolato sia rispetto al totale dei casi cumulati alla data di riferimento che considerando gli scarti tra i casi a tutti i passi temporali.

<sup>11</sup> Calcolato come:  $Errore\ 1 = \frac{casi\ stimati(al\ 30\ settembre\ 2011) - casi\ reali(al\ 30\ settembre\ 2011)}{casi\ reali(al\ 30\ settembre\ 2011)} \cdot 100$

<sup>12</sup> Calcolato come:  $Errore\ 2 = \sum (casi\ stimati - casi\ osservati)^2$ , calcolati per ogni nodo e per tutti i passi temporali.

## CONCLUSIONI

Introducendo la precipitazione come variabile nel modello epidemiologico si riesce a descrivere meglio il reale andamento dell'epidemia di colera divampata ad Haiti. È possibile quindi utilizzare le simulazioni come base per pianificare gli interventi da attuare per arginare la diffusione di questa malattia. In generale, è opportuno compiere delle previsioni a lungo termine da integrare con quelle a breve termine. In questo modo si possono pianificare dei programmi di intervento nel lungo periodo, che possono poi essere modificati in base alle esigenze che si vengono via via a verificare con interventi di brevissimo periodo.

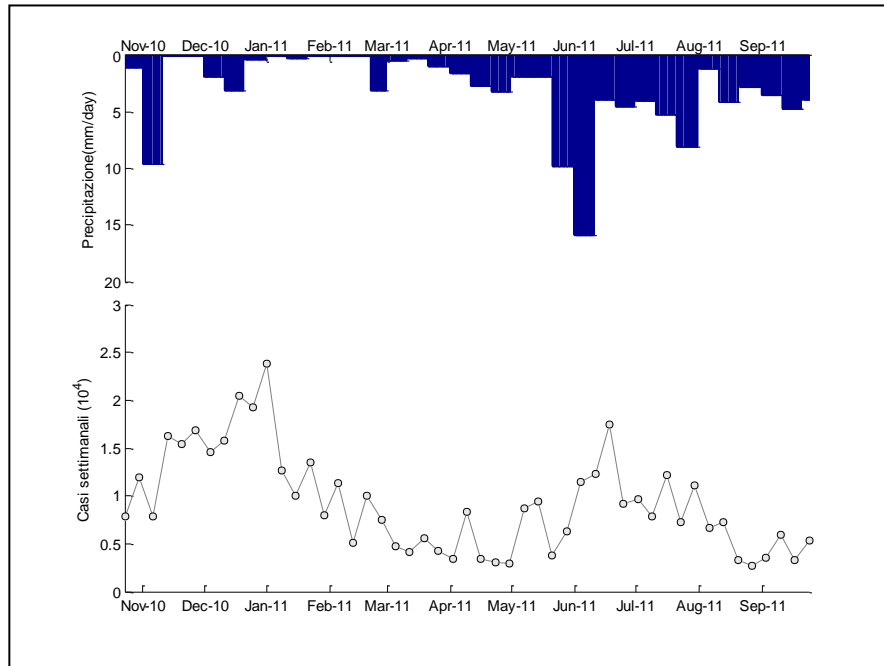
Le previsioni a lungo termine sono necessarie per pianificare strategie impegnative sia da un punto di vista temporale che economico. Fanno parte di questa categoria interventi atti a costruire infrastrutture idriche adeguate, a migliorare le condizioni igienico-sanitarie e a sviluppare programmi di sensibilizzazione della popolazione. Il gran numero di scenari generati consente di valutare la linea di azione migliore. Tutte le simulazioni concordano nel prevedere una prima risorgenza dell'epidemia di colera ad Haiti intorno a settembre 2012 e una seconda, più acuta, a settembre 2013. Si nota quindi l'andamento stagionale dell'epidemia che segue i periodi piovosi. Il metodo migliore per generare la precipitazione da introdurre nel modello epidemiologico è quello che utilizza le statistiche di precipitazione (media, 25-esimo percentile e 75-esimo percentile). Infatti, nonostante sia il metodo più semplice è comunque da preferire rispetto agli altri poiché offre una maggiore variabilità di scenari. Per generare un più ampio spettro di scenari anche con gli altri metodi, si dovrebbero apportare alcune migliorie, per esempio sarebbe opportuno raffinare la scala temporale del generatore puntuale di precipitazione poissoniano, oppure introdurre un termine per considerare esplicitamente la correlazione spaziale dei bacini.

Per quanto riguarda le previsioni a breve termine, un modello robusto di generazione della precipitazione è il modello ARX, che considera esplicitamente le correlazioni temporali e spaziali. L'utilizzo, da parte dei modelli ARX, di tutte le informazioni disponibili li rende inoltre particolarmente adatti all'utilizzo in una

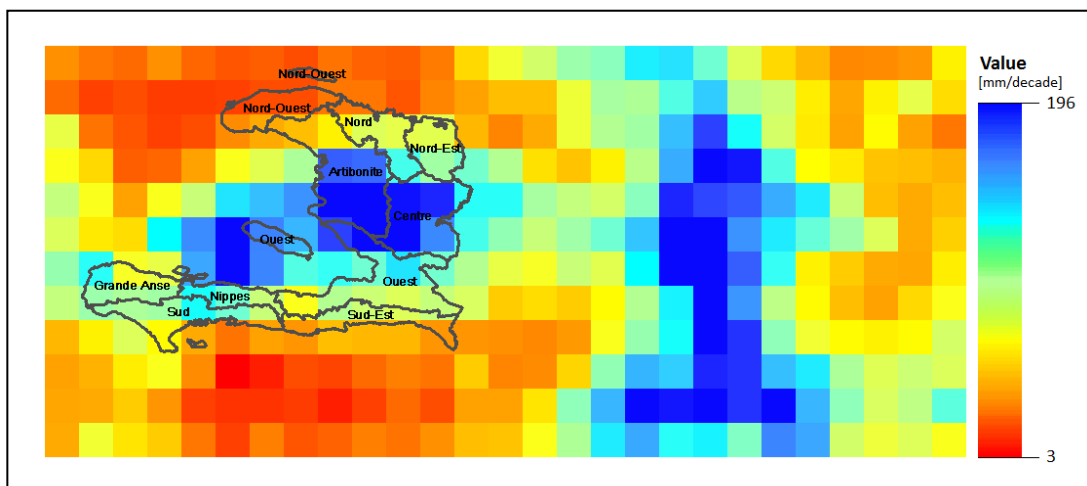
politica di controllo in linea dell'epidemia. Alcuni interventi che possono essere implementati anche a brevissimo termine sono, ad esempio, l'approvvigionamento di acqua potabile o la spedizione di kit igienico-sanitari nelle zone che si prevede saranno più colpite.

Infine, si confrontano i dati reali aggiornati al 20 gennaio 2012 (**Figura 20**), con quelli simulati con l'anno di precipitazione medio. È possibile osservare un aumento delle precipitazioni reali, che da settembre 2011 rientrano nell'intervallo tra il 25-esimo percentile e il 75-esimo percentile di precipitazione. In particolare, dall'inizio della stagione secca tipica di Haiti (dicembre 2011) l'andamento della precipitazione reale e simulata sono molto simili (**Figura 20c**). Di conseguenza anche gli andamenti dei casi settimanali reali e simulati corrispondono (**Figura 20a**) e quindi anche la differenza tra i casi cumulati reali e simulati diminuisce (**Figura 20b**). Le previsioni effettuate riescono quindi a riprodurre il reale andamento dell'epidemia di colera tranne in situazioni particolari, come può essere la mancanza della seconda stagione piovosa avvenuta nel 2011.

Effettuando delle previsioni affidabili sulla diffusione del colera e definendo un gran numero di possibili scenari di precipitazione e di conseguenza anche epidemiologici, i modelli analizzati in questa tesi, rappresentano quindi un efficace strumento di supporto decisionale durante tutta l'evoluzione dell'epidemia.

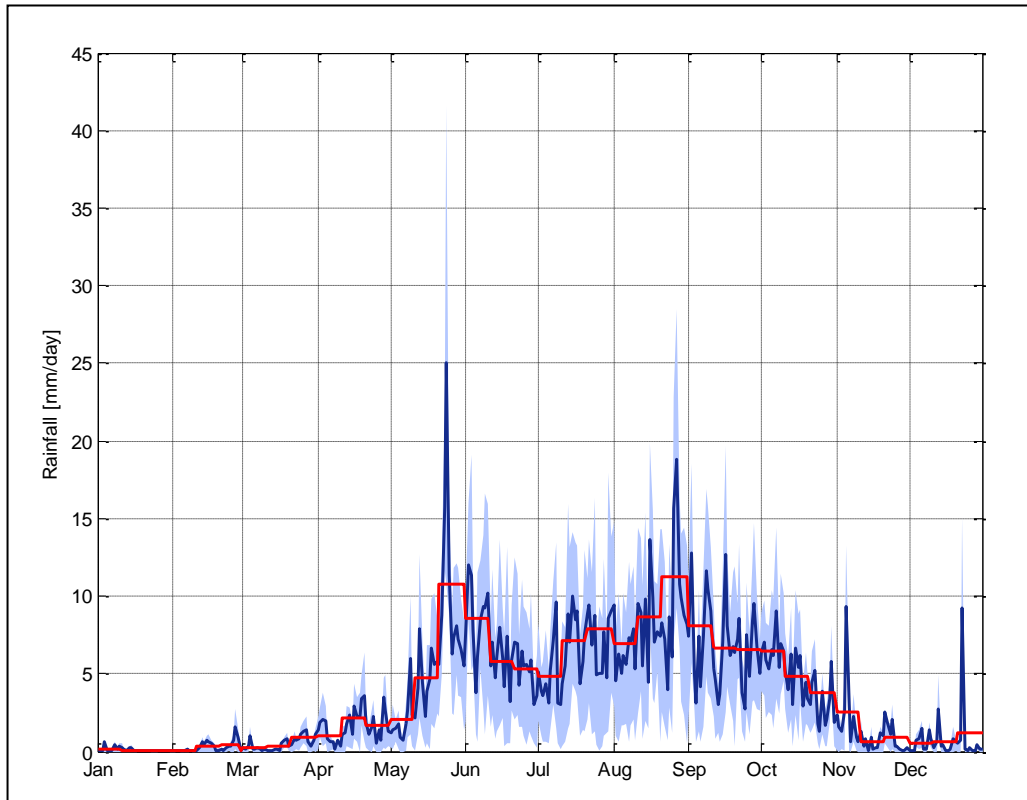


**Figura 1:** Nel grafico in alto sono riportate le precipitazioni dal 23 ottobre 2010 fino al 30 settembre 2011. Nel grafico in basso sono rappresentati i casi di colera settimanali registrati. È possibile notare una corrispondenza tra l’inizio del periodo piovoso a maggio/giugno e il secondo picco di casi di colera.

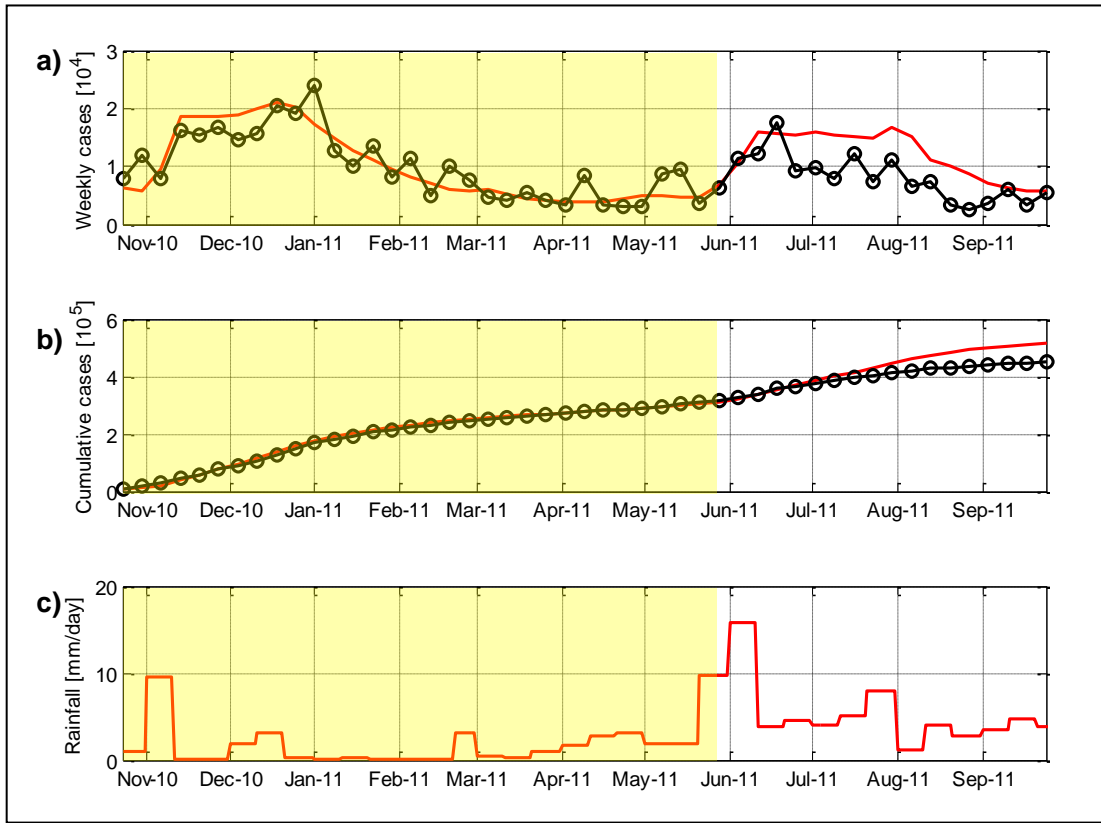


**Figura 2:** Rappresentazione dei dati di precipitazione decadali forniti dall’*International Research Institute for Climate and Society*. Si riporta soltanto Haiti. I dati sono relativi alla terza decade di luglio 2011. Ogni immagine è composta da 12x27 celle, ognuna delle quali di lato di 0.25 gradi (circa 28 km).





**Figura 3:** Media dell'andamento annuale della precipitazione di Haiti. In blu è riportato l'andamento della precipitazione media giornaliera, in rosso quello della precipitazione media decadale, mentre in azzurro è rappresentato l'intervallo di precipitazione giornaliera tra il 25-esimo percentile e il 75-esimo percentile.



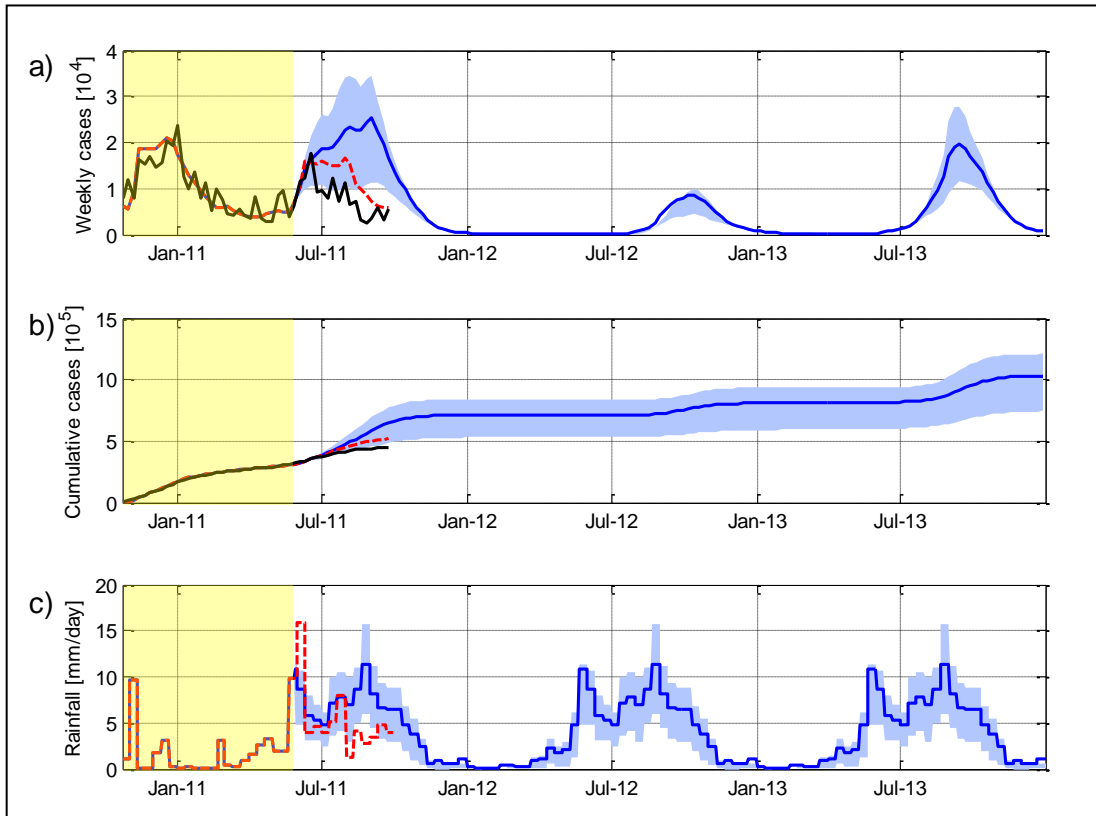
**Figura 4:** Calibrazione e validazione del modello epidemiologico (Capitolo 2, Modello 2.4).

a) Andamento dei casi settimanali di colera. In nero sono indicati i casi osservati, mentre in rosso quelli simulati con la precipitazione reale per il periodo dal 23 ottobre 2010 al 30 settembre 2011.

b) Andamento dei casi di colera cumulati. In nero sono indicati i casi osservati, mentre in rosso quelli simulati con la precipitazione reale.

c) Andamento delle precipitazioni reali mediate su tutta Haiti.

Lo sfondo giallo indica il periodo della calibrazione del modello, il bianco quello della validazione.



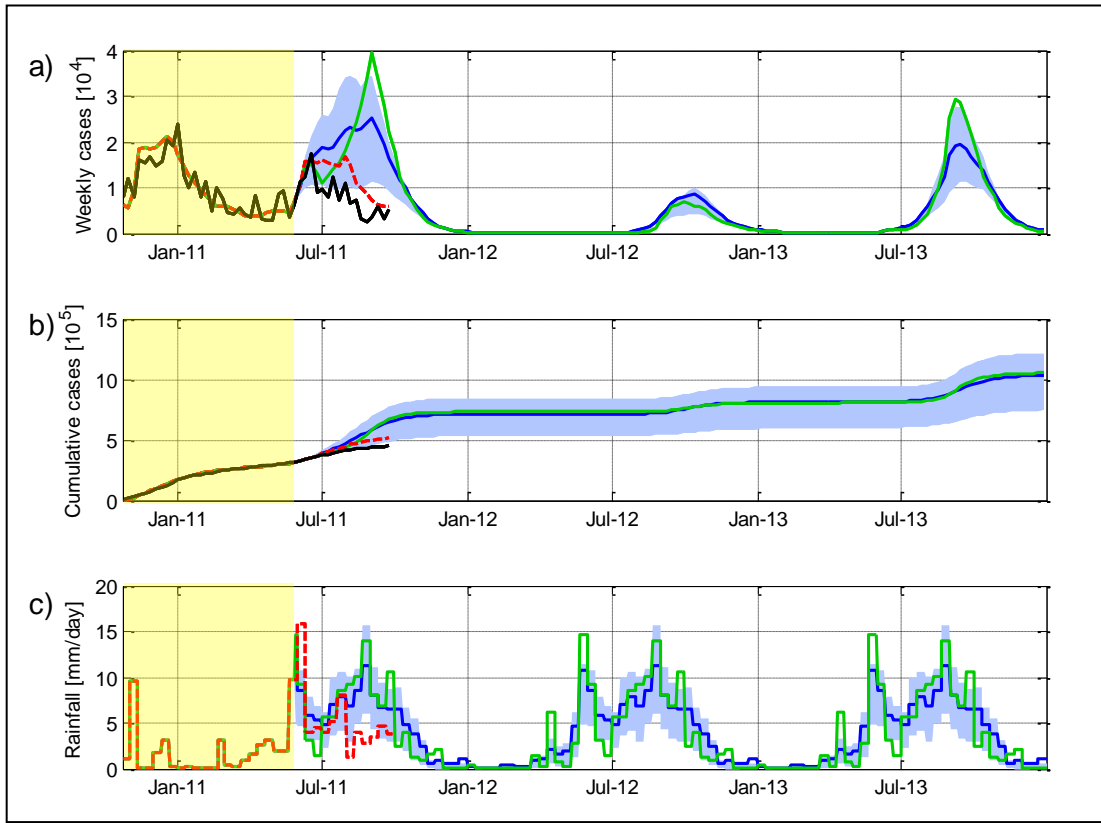
**Figura 5:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) con l'anno di precipitazione medio.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, con una linea tratteggiata rossa quelli simulati utilizzando i dati di precipitazione reali e in blu quelli che utilizzano l'anno medio di precipitazione. Infine, la banda azzurra rappresenta l'intervallo dei casi simulati utilizzando il 25-esimo percentile delle precipitazioni per ogni nodo e il 75-esimo percentile.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In rosso è rappresentato l'andamento della pioggia decadale secondo i dati della *NASA-JAXA's Tropical Rainfall Measuring Mission*, mentre in blu quello delle precipitazioni medie, usate per la simulazione. La banda azzurra rappresenta l'intervallo tra la media su Haiti del primo quartile e del terzo quartile delle precipitazioni di ogni nodo.

Lo sfondo giallo indica il periodo della calibrazione.



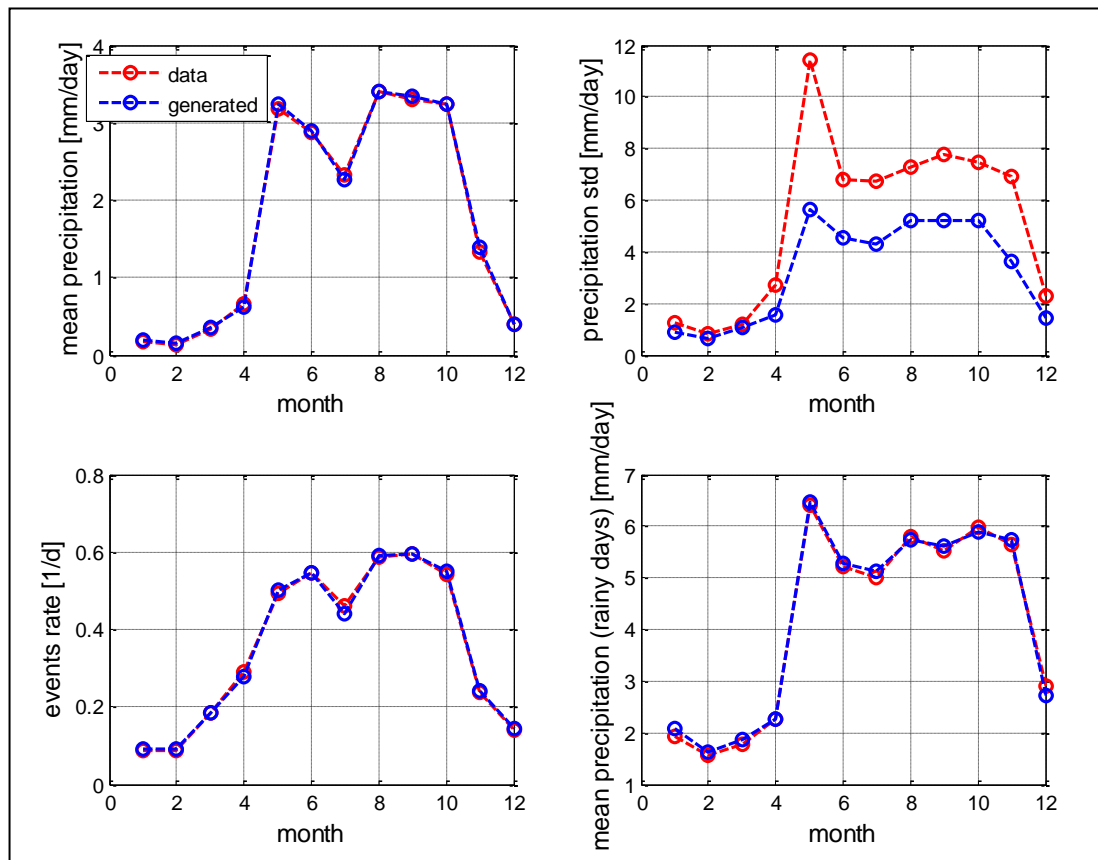
**Figura 6:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con la precipitazione dell'anno tipico.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera reali, con una linea tratteggiata rossa quelli simulati utilizzando i dati di precipitazione reali, in verde i casi ottenuti con la precipitazione dell'anno tipico e in blu quelli che utilizzano l'anno medio di precipitazione. Infine, la banda azzurra rappresenta l'intervallo dei casi simulati utilizzando il 25-esimo percentile delle precipitazioni per ogni nodo e il 75-esimo percentile.

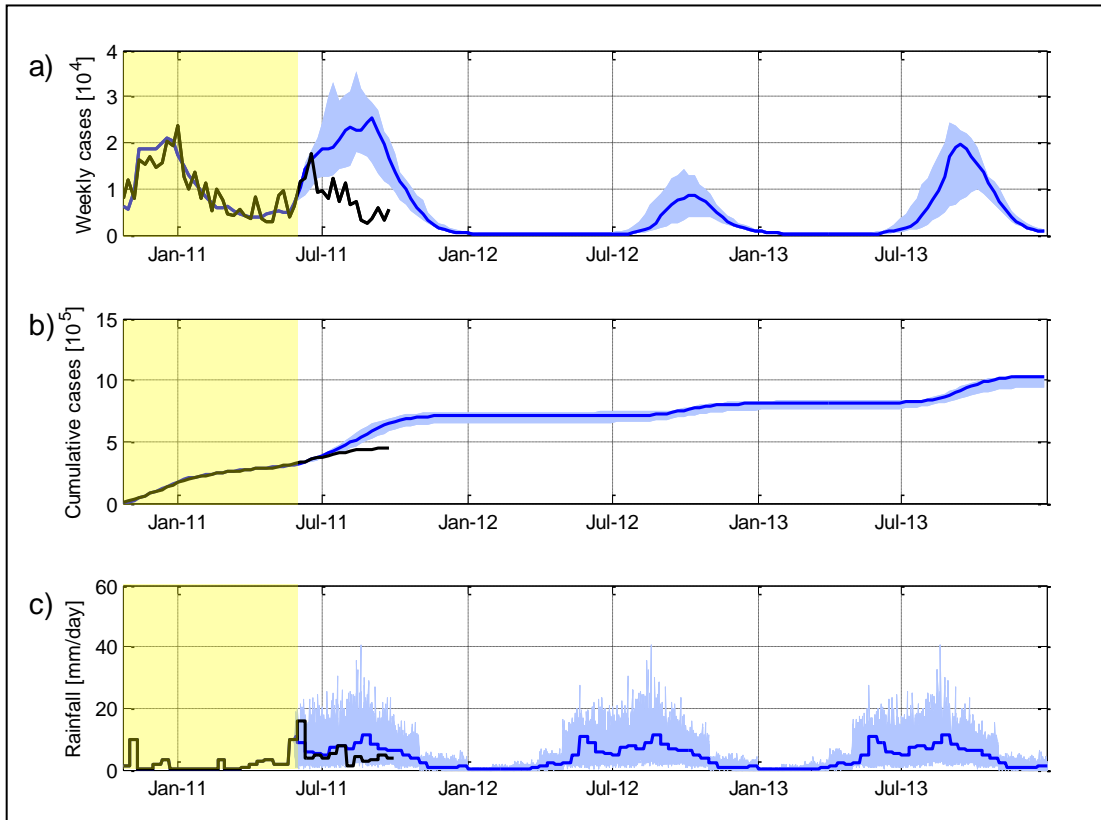
b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In rosso è rappresentato l'andamento della pioggia decadale secondo i dati registrati dalla *NASA-JAXA's Tropical Rainfall Measuring Mission*, in verde è riportato l'andamento delle precipitazioni dell'anno tipico (che corrisponde al 2002) e in blu quello delle precipitazioni medie. La banda azzurra rappresenta l'intervallo tra la media su Haiti del primo quartile e del terzo quartile delle precipitazioni di ogni nodo.

Lo sfondo giallo indica il periodo della calibrazione.



**Figura 7:** Statistiche calcolate sui dati generati con il generatore puntuale di precipitazione di Poisson per il distretto di Nord-Ouest. In rosso sono rappresentate le statistiche reali, mentre in blu quelle della precipitazione generata dal generatore.



**Figura 8:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei dipartimenti disaggregati.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, mentre in blu sono indicati quelli simulati con l'anno di precipitazione media. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dai dipartimenti disaggregati.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In nero sono raffigurati i dati reali, mentre in blu sono indicate le precipitazioni medie. La fascia azzurra rappresenta l'intervallo delle precipitazioni generate con il generatore puntuale di precipitazione di Poisson mediate sui nodi per ognuno dei cento anni generati (quindi si è già effettuato il *downscaling*) a partire dai dipartimenti disaggregati.

Lo sfondo giallo indica il periodo della calibrazione.

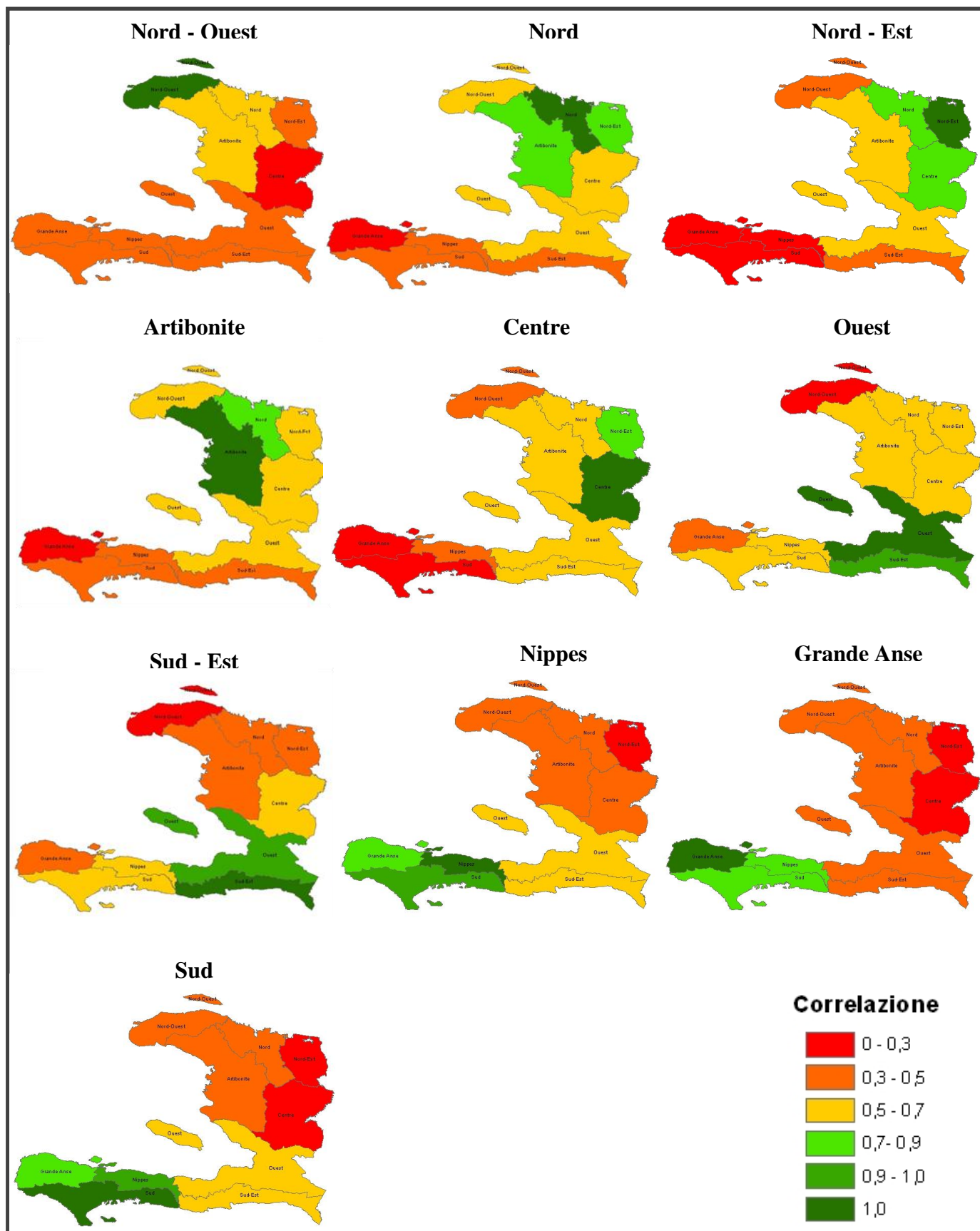
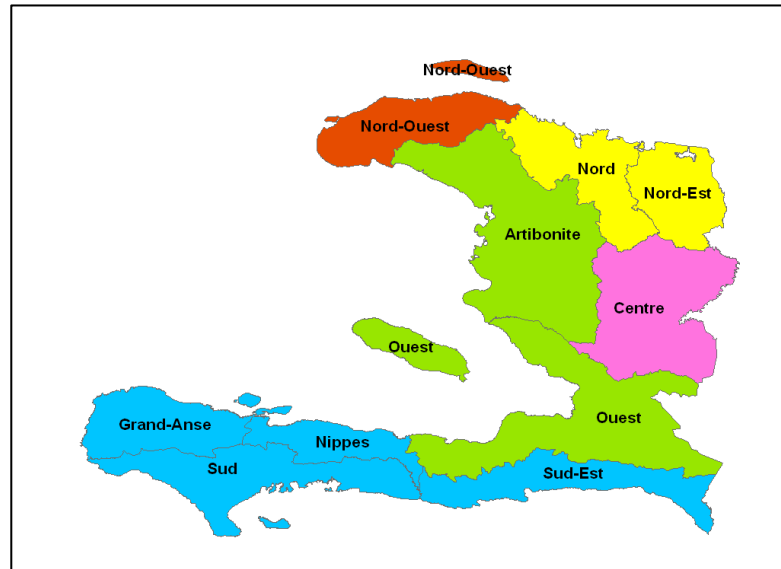
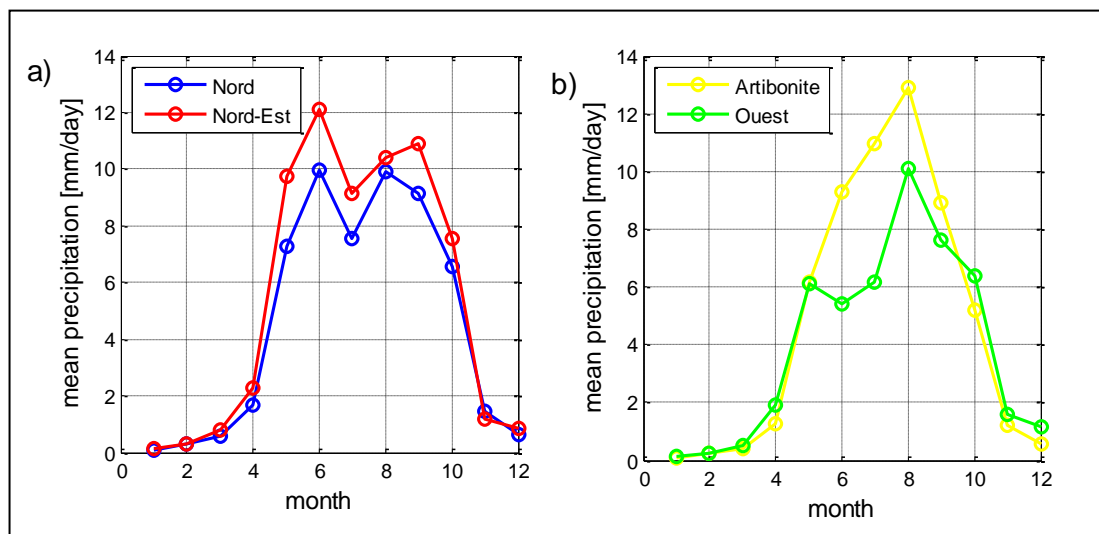


Figura 9: Mappe di correlazione tra la precipitazione di un dipartimento (in verde scuro) e quella dei distretti circostanti.

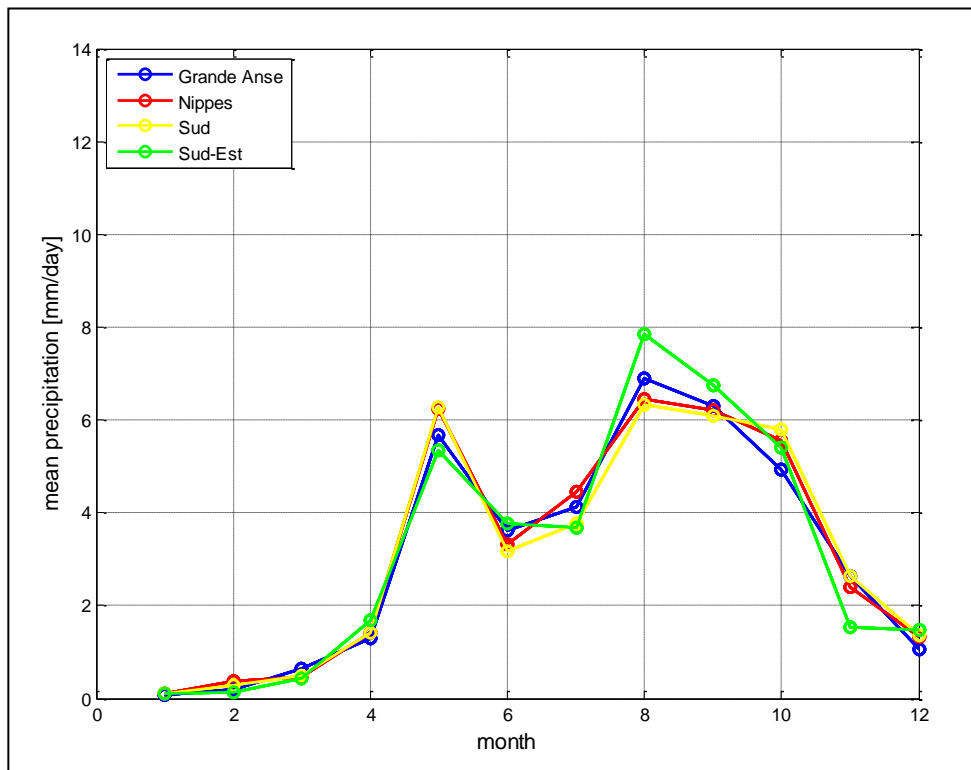


**Figura 10:** Suddivisione di Haiti in zone con precipitazioni comparabili.

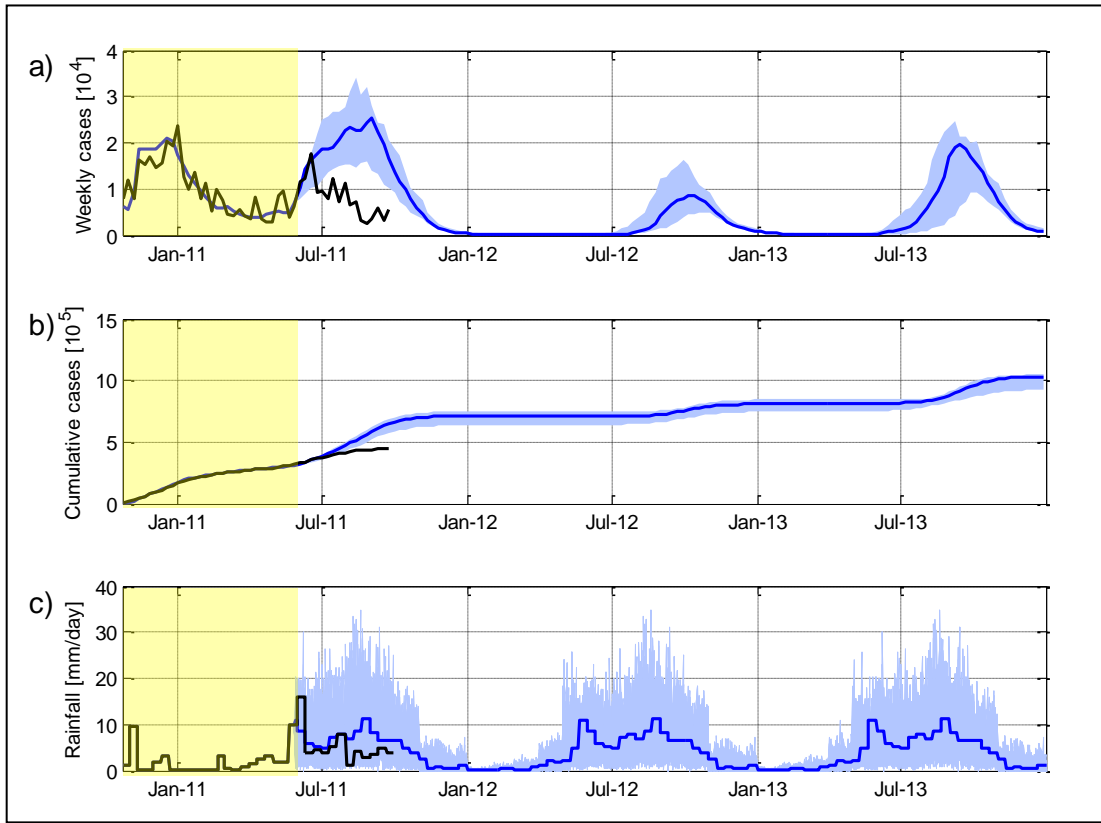


**Figura 11:** Andamento delle precipitazioni medie mensili di:  
 c) Distretti di Nord (in blu) e Nord-Est (in rosso)  
 d) Distretti di Artibonite (in giallo) e di Ouest (in verde)





**Figura 12:** Andamento delle precipitazioni medie mensili dei distretti di Grande Anse (in blu), Nippes (in rosso), Sud (in giallo) e Sud-Est (in verde).



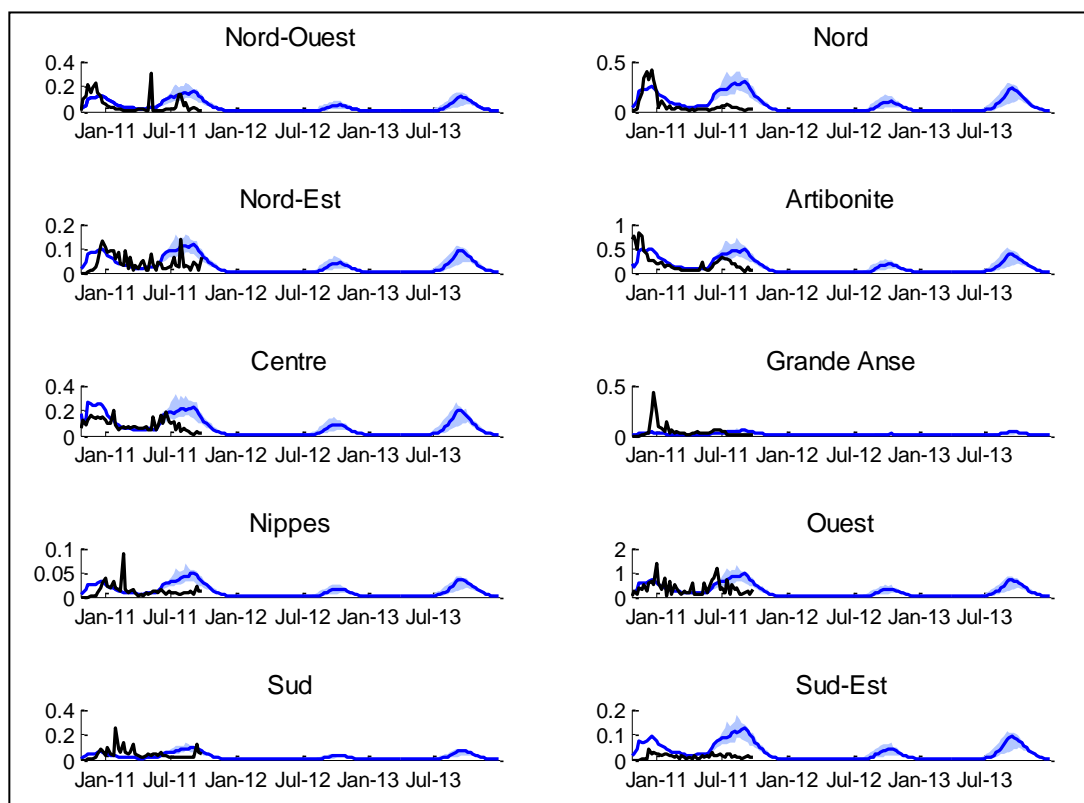
**Figura 13:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei dipartimenti aggregati.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, mentre in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dai dipartimenti aggregati.

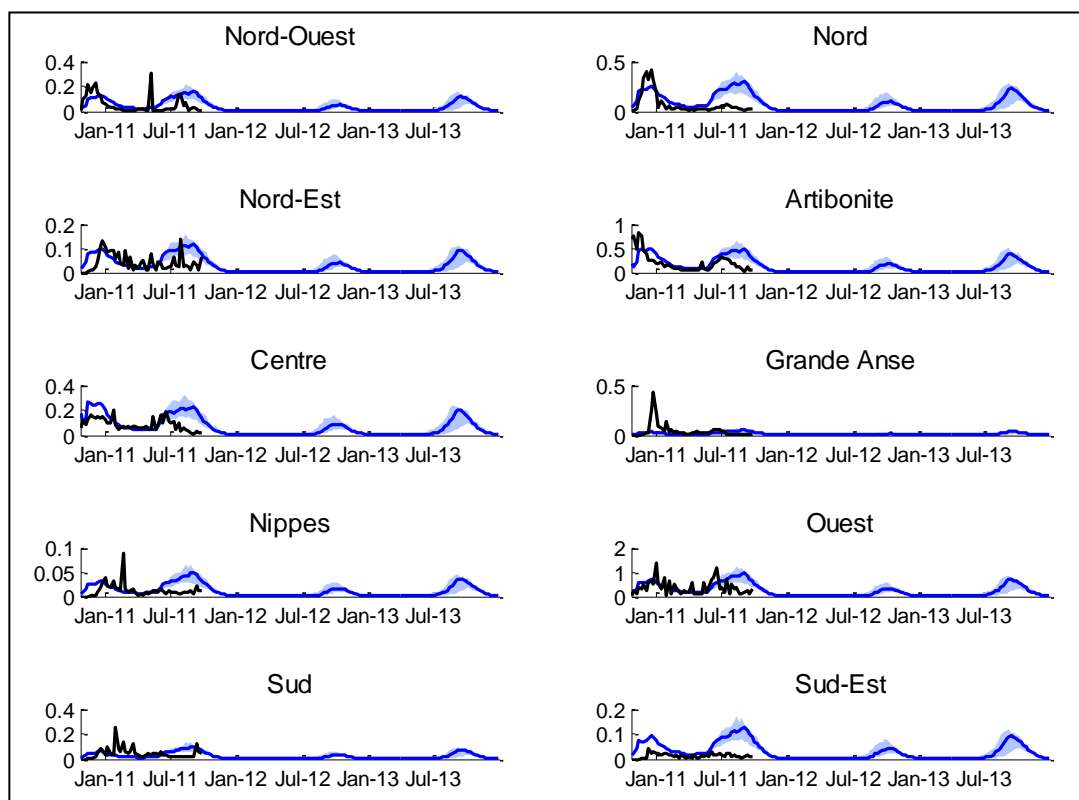
b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In nero sono raffigurati i dati reali, mentre in blu sono indicate le precipitazioni medie. La fascia azzurra rappresenta l'intervallo delle precipitazioni giornaliere (mediate sui nodi per ognuno dei cento anni generati) generate con il generatore puntuale di Poisson a partire dai dipartimenti aggregati.

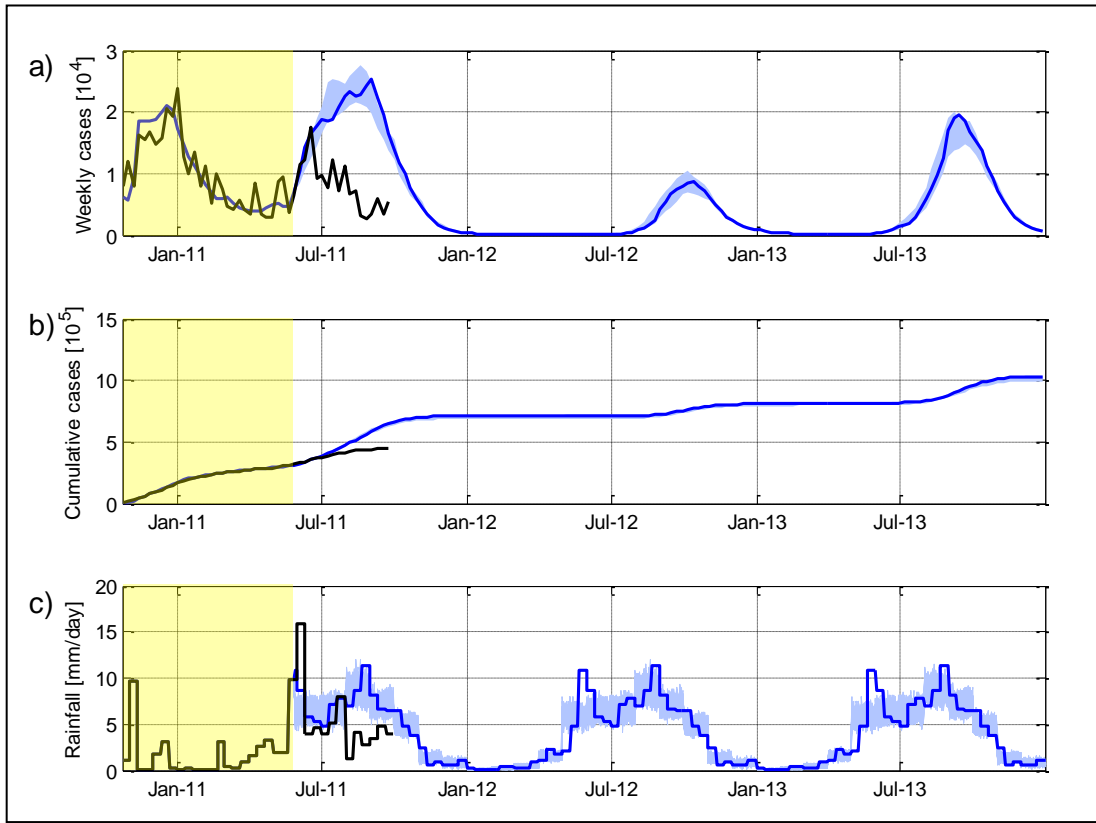
Lo sfondo giallo indica il periodo della calibrazione.



**Figura 14:** Casi di colera settimanali per ogni distretto simulati con il modello epidemiologico (Capitolo 2, Modello 2.4) che utilizza come ingresso la precipitazione generata con il generatore di precipitazione puntuale di Poisson partendo dalla precipitazione giornaliera dei distretti disaggregati. Sull'asse delle ordinate i casi sono riportati divisi per  $10^4$ . In nero sono riportati i casi di colera registrati, mentre in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dalla precipitazione dei dipartimenti disaggregati.



**Figura 15:** Casi di colera settimanali per ogni distretto simulati con il modello epidemiologico (Capitolo 2, Modello 2.4) che utilizza come ingresso la precipitazione generata con il generatore puntuale di precipitazione di Poisson partendo dalla precipitazione giornaliera dei distretti aggregati. Sull'asse delle ordinate i casi sono riportati divisi per  $10^4$ . In nero sono riportati i casi di colera registrati, mentre in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dai dipartimenti aggregati.



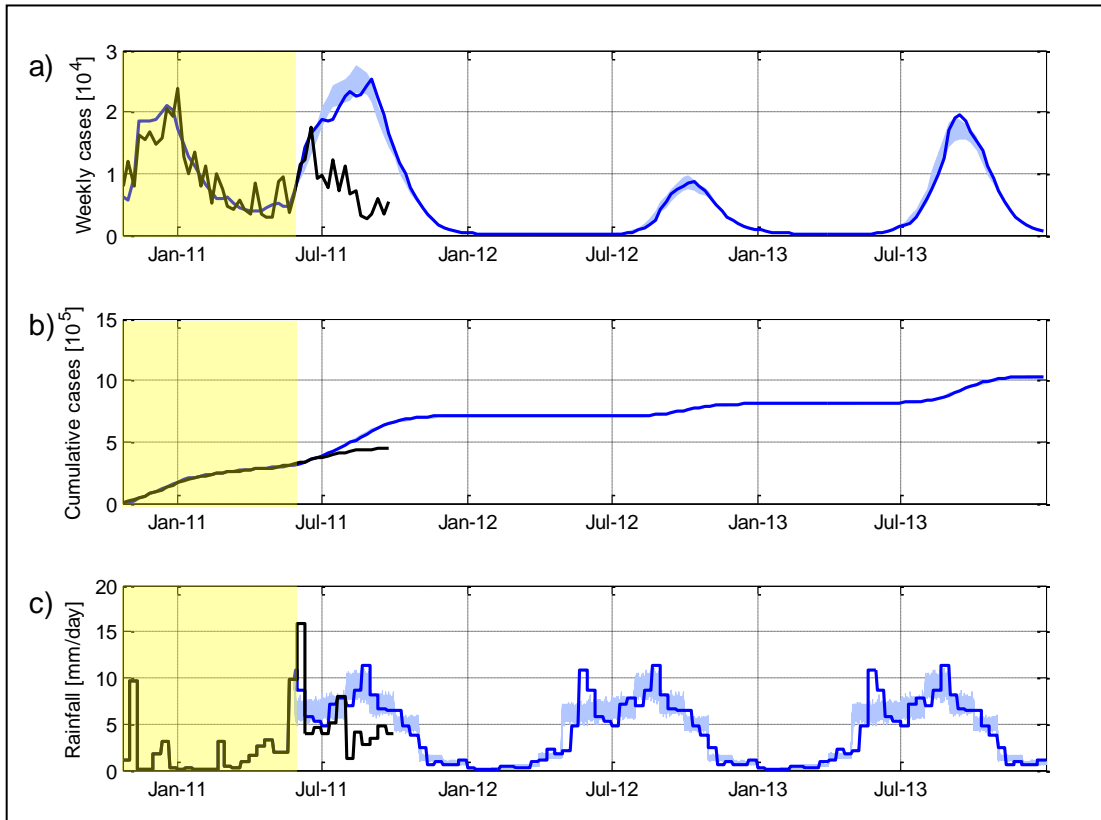
**Figura 16:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei nodi.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, mentre in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei nodi.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In nero sono raffigurati i dati reali, mentre in blu sono indicate le precipitazioni medie. La fascia azzurra rappresenta l'intervallo delle precipitazioni giornaliere (mediate sui nodi per ognuno dei cento anni generati) generate con il generatore puntuale di Poisson a partire dalla precipitazione giornaliera dei nodi.

Lo sfondo giallo indica il periodo della calibrazione.



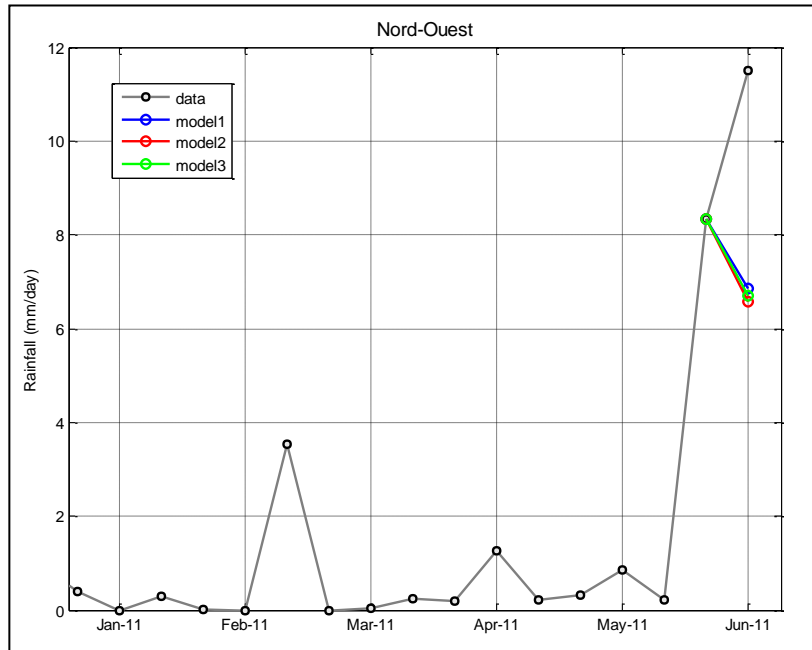
**Figura 17:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con le precipitazioni generate dal generatore puntuale di Poisson a partire dalla precipitazione decadale dei nodi.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati e in blu sono indicati quelli simulati con l'anno di precipitazione medio. La fascia azzurra rappresenta l'intervallo tra i valori minimi e massimi degli scenari ottenuti con le precipitazioni generate con il generatore puntuale di Poisson a partire dalla precipitazione decadale dei nodi.

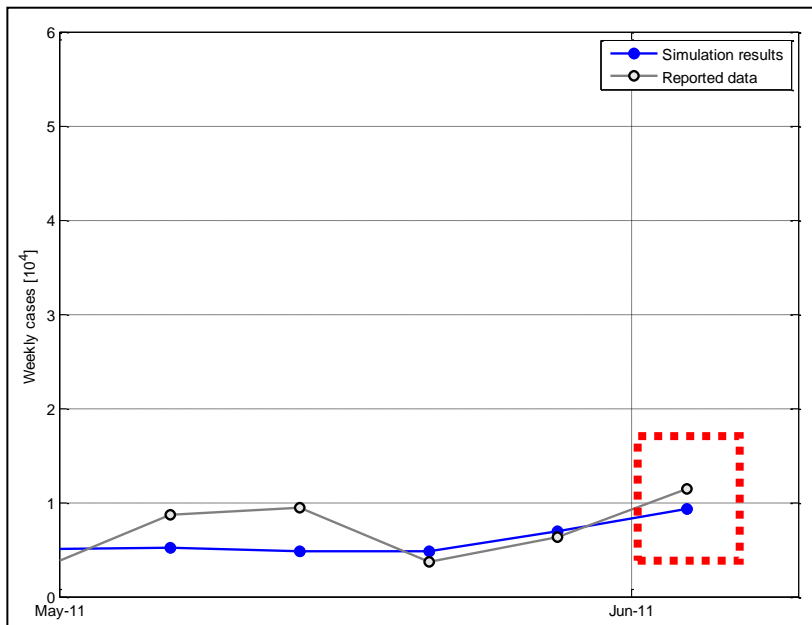
b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In nero sono raffigurati i dati reali, mentre in blu sono indicate le precipitazioni medie. La fascia azzurra rappresenta l'intervallo delle precipitazioni giornaliere (mediate sui nodi per ognuno dei cento anni generati) generate con il generatore puntuale di Poisson a partire dalla precipitazione decadale dei nodi.

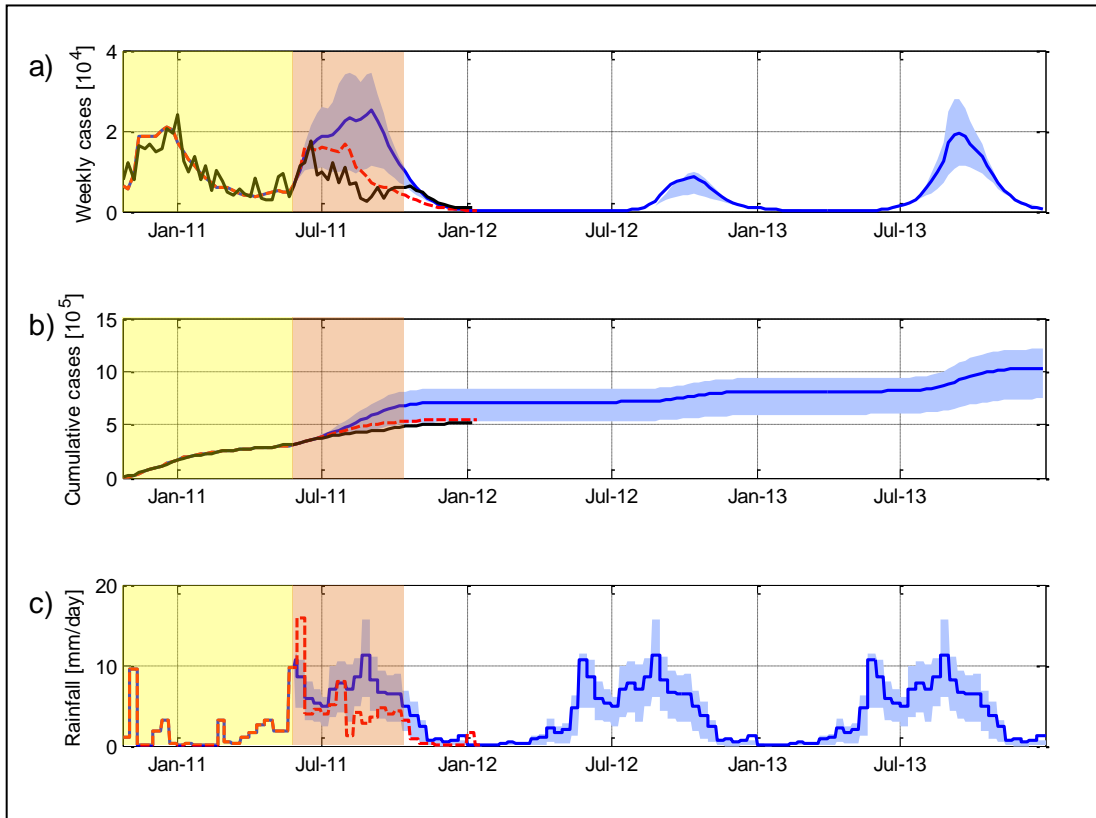
Lo sfondo giallo indica il periodo della calibrazione.



**Figura 18:** Precipitazione simulata con i modelli ARX per il distretto Nord-Ovest. Si noti che ogni dato rappresenta quanto piove in un giorno di una data decade. La previsione si riferisce alla prima decade di giugno 2011.



**Figura 19:** Zoom temporale sulla simulazione dei casi di colera simulati con il modello epidemico (Capitolo 2, Modello 2.4) basandosi sulla precipitazione generata utilizzando i modelli ARX. In blu sono riportati i dati generati, mentre con i pallini grigi i dati reali. Il dato dei casi di colera simulato con la precipitazione generata dai modelli ARX, che si riferisce al 4 giugno 2011, è contornato con un tratteggio rosso.



**Figura 20:** Simulazione del modello epidemiologico (Capitolo 2, Modello 2.4) eseguita con l'anno di precipitazione medio. Confronto con i dati reali aggiornati al 20 gennaio 2012.

a) Casi settimanali di colera sommati su tutto il territorio di Haiti. In nero sono riportati i casi di colera registrati, con una linea tratteggiata rossa quelli simulati utilizzando i dati di precipitazione reali e in blu quelli che utilizzano l'anno medio di precipitazione. Infine, la banda azzurra rappresenta l'intervallo dei casi simulati utilizzando il 25-esimo percentile delle precipitazioni per ogni nodo e il 75-esimo percentile.

b) Casi cumulati di colera. I colori corrispondono al grafico (a), con l'unica differenza che i casi sono cumulati e non settimanali.

c) Andamento delle precipitazioni mediate su tutta Haiti. In rosso è rappresentato l'andamento della pioggia decadale secondo i dati della *NASA-JAXA's Tropical Rainfall Measuring Mission*, mentre in blu quello delle precipitazioni medie, usate per la simulazione. La banda azzurra rappresenta la media su Haiti del primo quartile e del terzo quartile delle precipitazioni di ogni nodo.

Lo sfondo giallo indica il periodo della calibrazione. Lo sfondo rosa indica il periodo dal 29 maggio 2011 al 30 settembre 2011.