

- BREVE PRESENTAZIONE:

Per questo progetto ho preso in esame le temperature massime registrate sul territorio italiano dal 1965 al 2022, utilizzando i dati forniti dal sito dell'Istituto Superiore per la Protezione e la Ricerca Ambientale. Ho raccolto e pulito i dati e tramite MySQL li ho racchiusi in un database suddiviso per Regioni. Ho quindi deciso di raggruppare le varie Regioni in tre macro-aree: Nord (Piemonte, Valle d'Aosta, Liguria, Lombardia, Veneto, Friuli-Venezia Giulia, Trentino-Alto Adige, Emilia Romagna), Centro (Lazio, Toscana, Marche, Umbria, Abruzzo) e Sud (Basilicata, Campania, Puglia, Molise, Calabria, Sardegna, Sicilia) e ho applicato a ciascuna di esse le metodologie considerate a lezione (in particolare l'analisi GEV e GPD), utilizzando Python e la libreria pyextremes.

Mentre la distribuzione GPD, nonostante numerosi tentativi, ha dimostrato di non adattarsi bene ai dati trattati, ho raccolto informazioni decisamente migliori tramite l'analisi GEV. In questo riassunto ho quindi deciso di presentare in maniera più compatta e fruibile i risultati relativi a quest'ultima.

Nello specifico per ogni area geografica ho riportato:

- Gli estremi ricavati col metodo Block Maxima (considerando intervalli di 365 giorni)
- I grafici Mean Residual Life e Paramater Stability
- Il grafico Peak Over Threshold (settato con intervalli di 30 giorni)
- Stima della temperatura massima raggiunta tra 3, 6, 10 e 15 anni, probabilità di superamento e periodo di ritorno
- I grafici sul periodo di ritorno, sulla densità di probabilità, Q-Q plot e P-P plot

In ogni macro-area le 10 temperature più alte registrate con intervalli di un anno si sono per la maggior parte verificate dal 1995 in poi, le prime 5 si sono sempre verificate dopo il 1995.

Analizzando il grafico POT (impostato con un intervallo di 30 giorni) ho potuto notare come in tutte le aree geografiche la soglia fissata sia stata superata più volte tra il 2000 e il 2022 che non tra il 1965 e il 1999. In particolare nel Nord Italia tra il 2000 e il 2022 in ben 17 occasioni l'estremo più alto registrato ha superato i 40 gradi, cosa che era avvenuta solo 4 volte tra il 1965 e il 1999. Per fissare la soglia del grafico POT ho fatto riferimento alla temperatura in cui il grafico Paramater Stability inizia a diventare meno stabile, ho inoltre osservato fino a quale temperatura i risultati del grafico Mean Residual Life sono collegabili con una retta.

Le stime delle temperature massime tra 3, 6, 10 e 15 anni fanno presupporre un aumento generalizzato delle temperature massime in tutta la penisola italiana. In particolare nel Centro e nel Sud Italia la stima per la temperatura massima registrata tra 15 anni è di almeno 2,5 gradi superiore a quella calcolata tra 3 anni, arrivando a quasi 3 gradi di differenza nel Centro Italia.

Analizzando la probabilità di superamento e il periodo di ritorno delle temperature più estreme registrate ogni anno, si può evincere come il Nord Italia faccia registrare eventi estremi meno frequenti e con temperature più basse. Il Centro Italia vede temperature più alte rispetto al Nord, con una frequenza simile per gli eventi estremi. Il Sud Italia non solo è la zona con le temperature massime più alte, ma anche quella dove gli eventi estremi fanno registrare una maggiore probabilità di superamento, segnalando una maggiore frequenza di eventi di caldo estremo rispetto alle altre aree del paese.

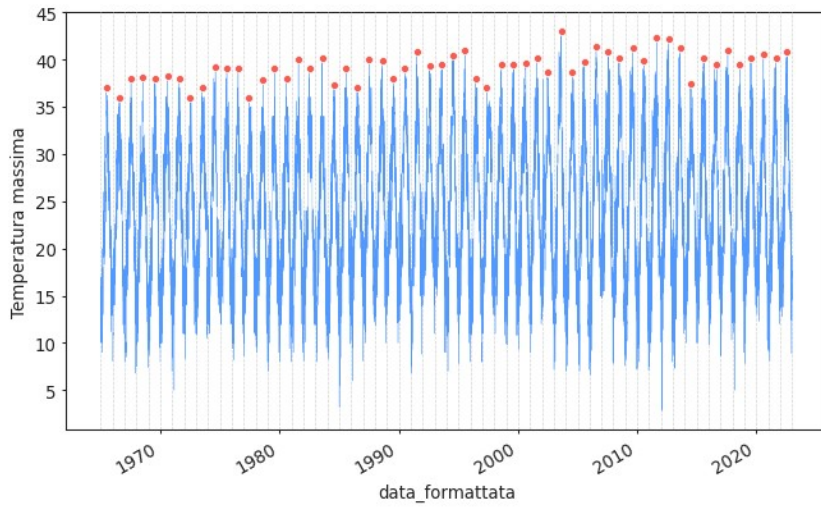
I grafici sul periodo di ritorno, sulla densità di probabilità, Q-Q plot e P-P plot fanno notare un grado di affidabilità delle previsioni molto buono, con plot dei valori predetti che ricalcano bene quelli effettivamente osservati e valori di R^2 molto vicini a 1.

In fondo ho riportato anche i grafici sul periodo di ritorno, sulla densità di probabilità, Q-Q plot e P-P plot relativi alla distribuzione GPD, per mostrare effettivamente quanto ben poco si adattasse ai dati trattati.

Il procedimento completo è consultabile sul file jupyter allegato alla mail.

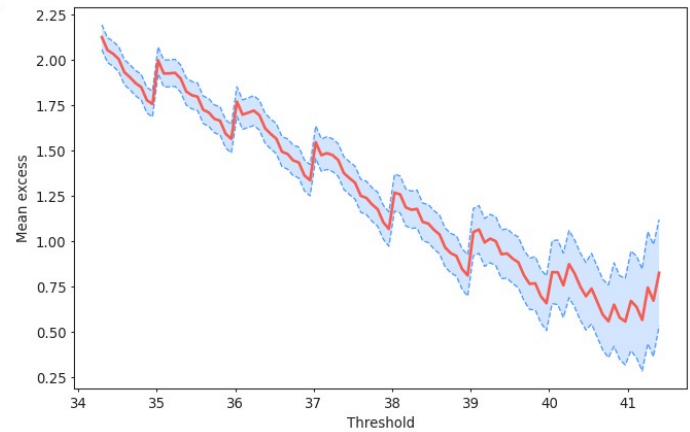
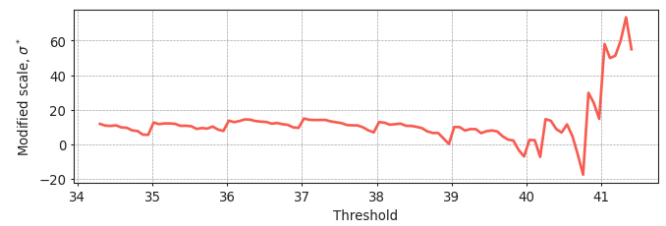
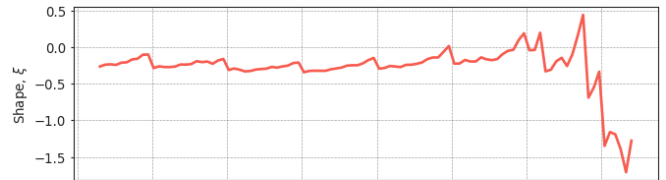
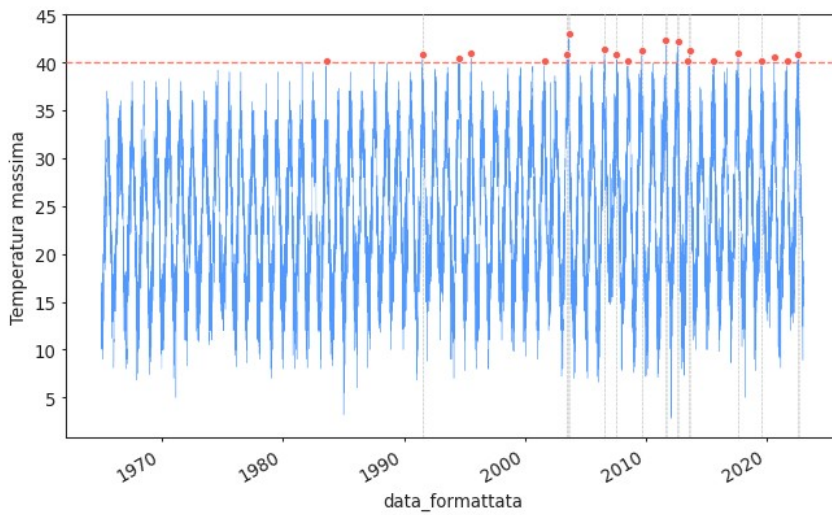
- RISULTATI

NORD ITALIA:



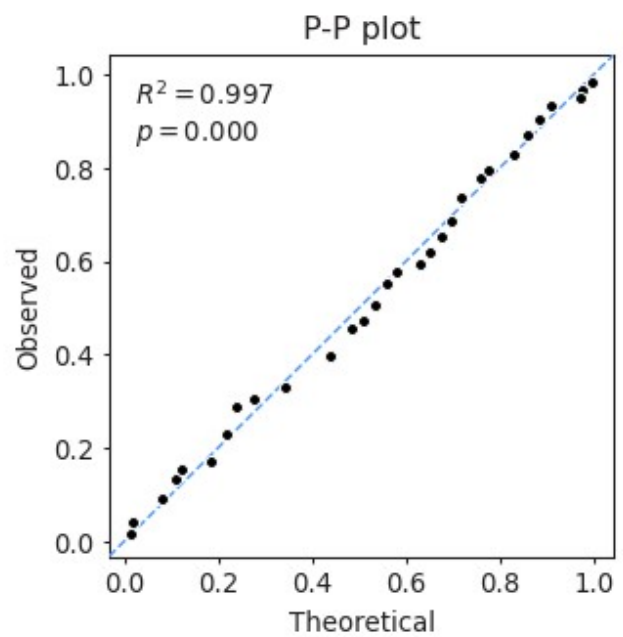
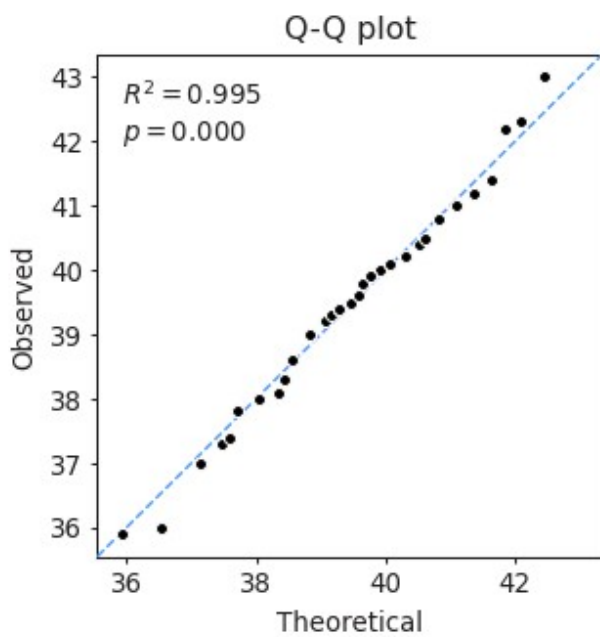
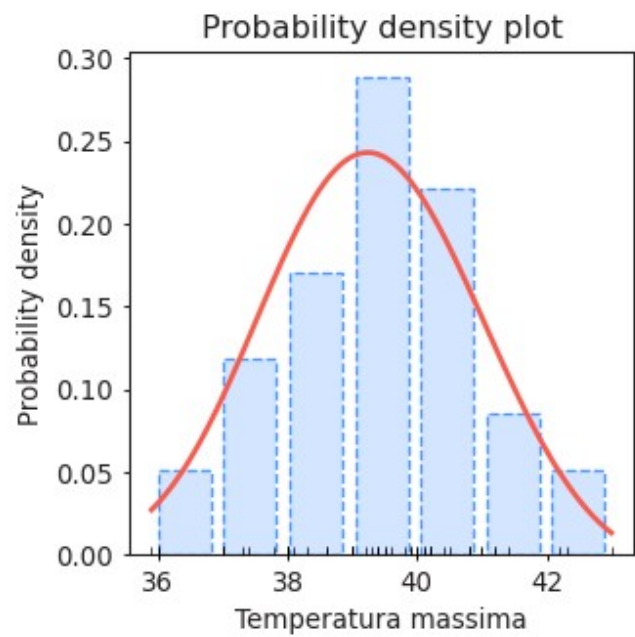
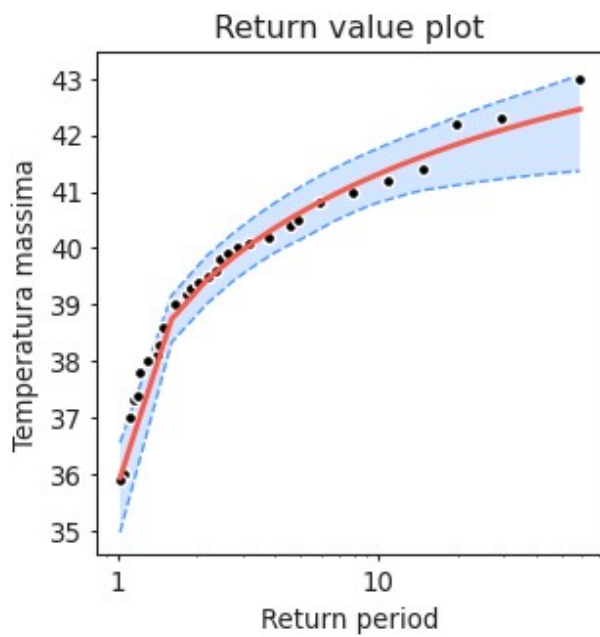
Le 10 temperature più alte registrate nel Nord Italia per ogni anno dal 1965 al 2022:

2003-08-11	43.0
2011-08-22	42.3
2012-08-22	42.2
2006-07-22	41.4
2009-08-20	41.2
2013-08-07	41.2
2017-08-04	41.0
1995-07-24	41.0
1991-07-14	40.8
2022-07-23	40.8

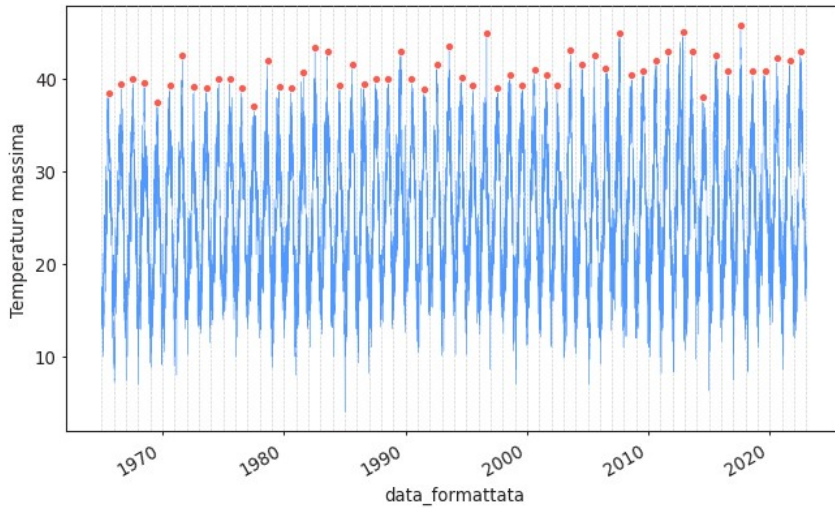


return period	return value	lower ci	upper ci
3.0	39.964286	39.527309	40.405913
6.0	40.829545	40.324900	41.265130
10.0	41.314908	40.744863	41.764049
15.0	41.639126	40.963852	42.085515

data	Temperatura massima	exceedance probability	return period
1988-07-23	45.0	0.016949	59.000000
2003-08-11	43.0	0.033898	29.500000
2011-08-22	42.3	0.050847	19.666667
2012-08-22	42.2	0.067797	14.750000
2006-07-22	41.4	0.084746	11.800000

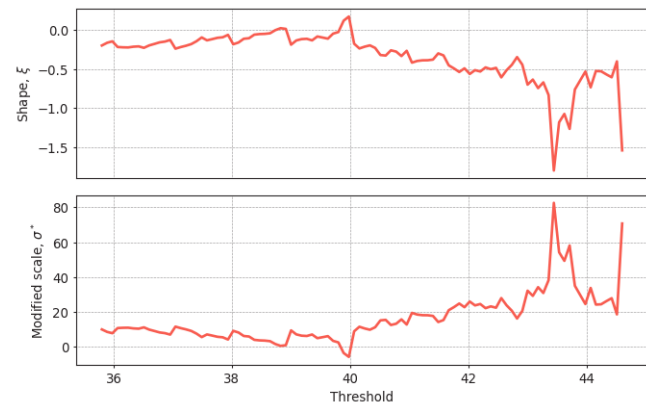
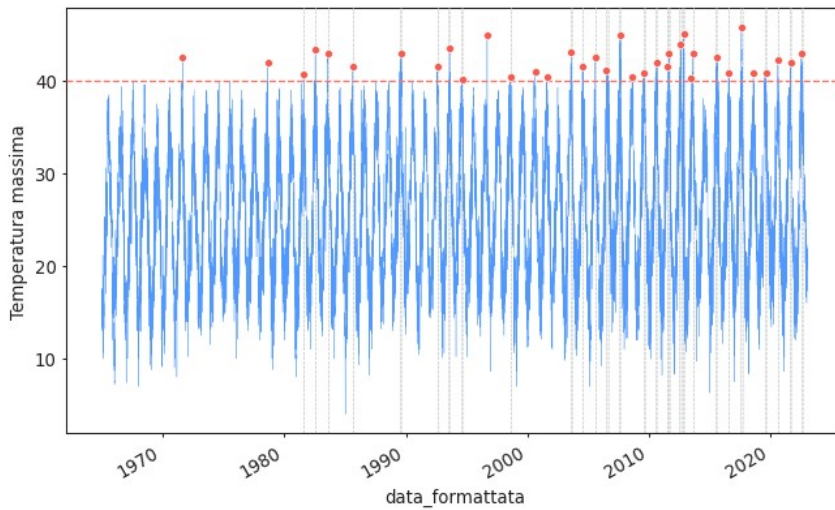


CENTRO ITALIA:



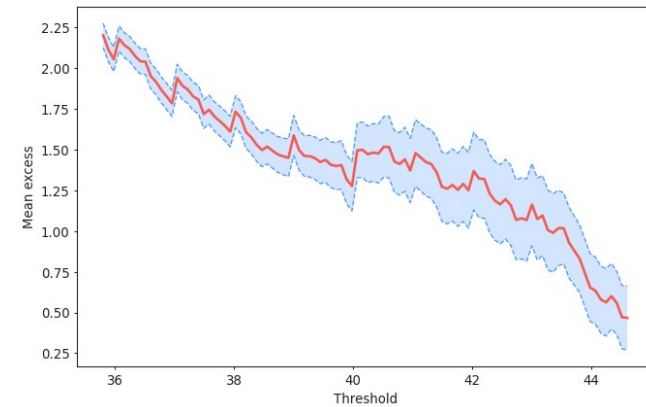
Le 10 temperature più alte registrate nel Centro Italia per ogni anno dal 1965 al 2022:

2017-08-04	45.8
2012-11-02	45.1
2007-08-30	45.0
1996-09-08	45.0
2012-08-07	44.0
1993-08-05	43.5
1982-07-22	43.4
2003-08-07	43.1
1983-07-27	43.0
1989-08-08	43.0

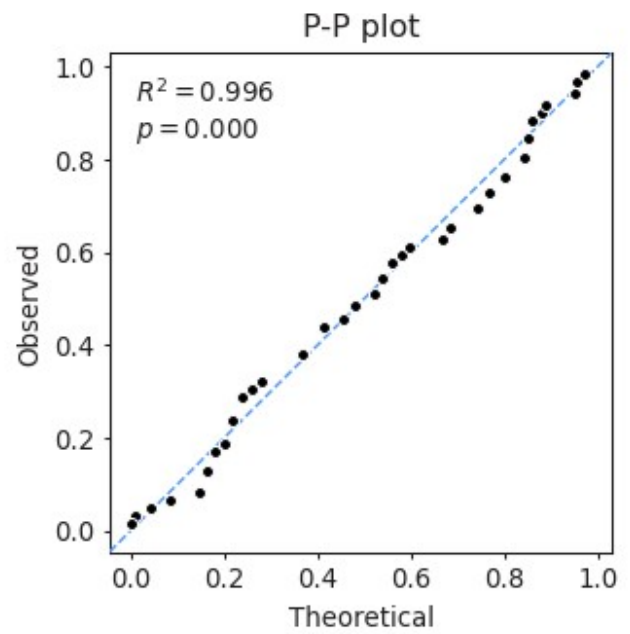
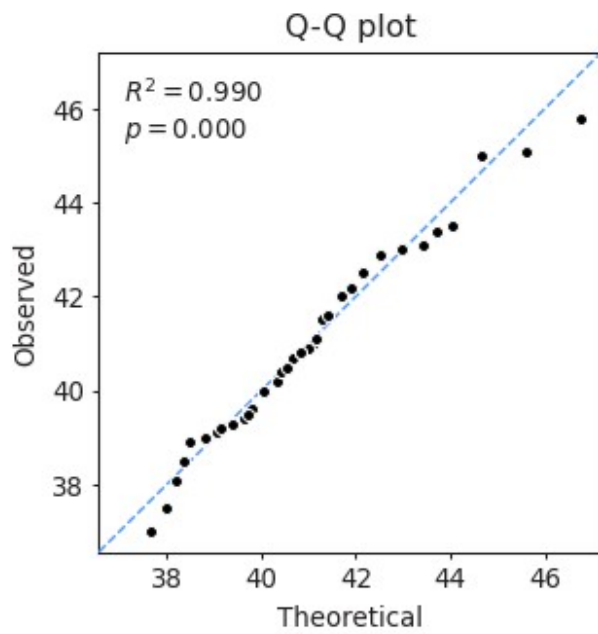
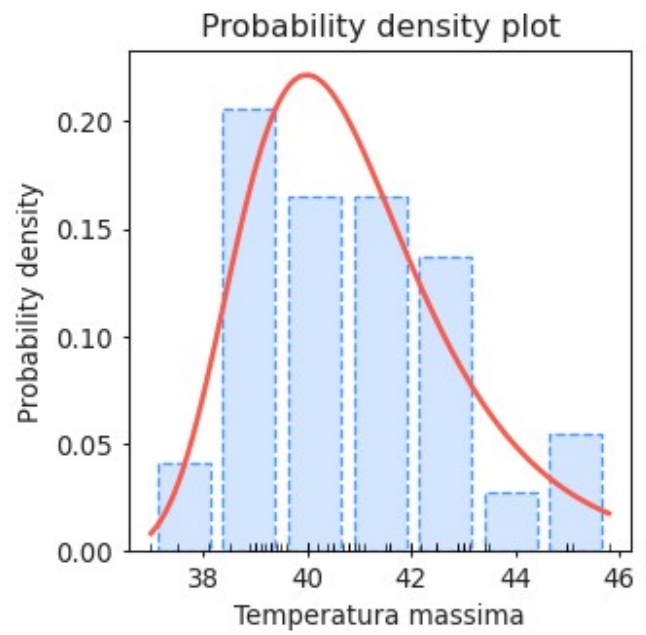
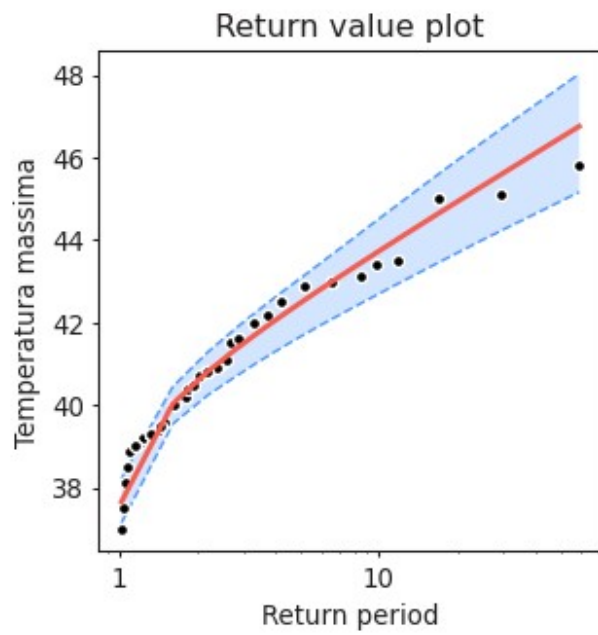


(threshold=42)

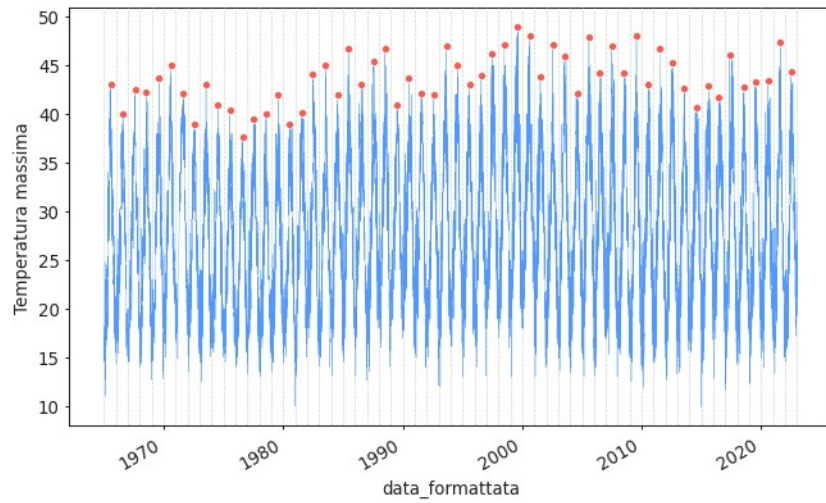
return period	return value	lower ci	upper ci
3.0	41.494540	40.942331	42.059862
6.0	42.822404	42.042408	43.552036
10.0	43.733468	42.745021	44.540974
15.0	44.436863	43.317048	45.344122



data	Temperatura massima	exceedance probability	return period
2017-08-04	45.8	0.016949	59.000000
2012-11-02	45.1	0.033898	29.500000
1996-09-08	45.0	0.059322	16.857143
2007-08-30	45.0	0.059322	16.857143
1993-08-05	43.5	0.084746	11.800000

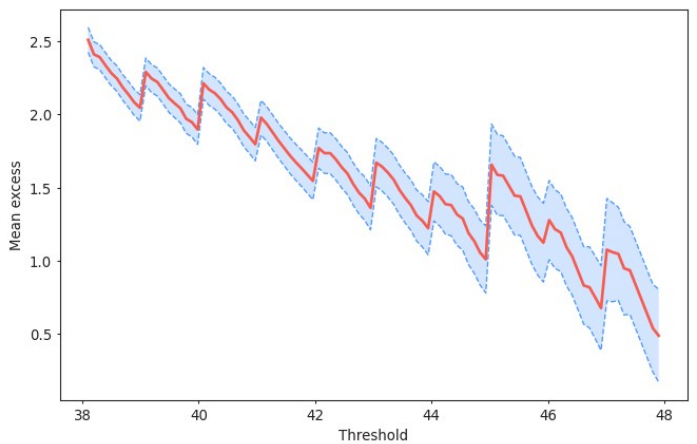
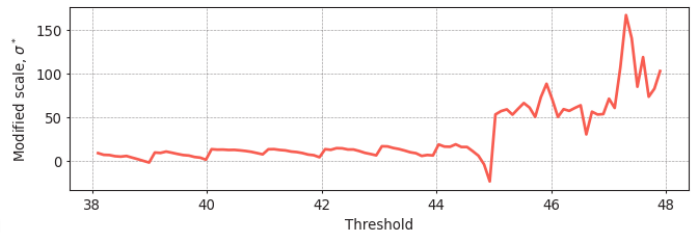
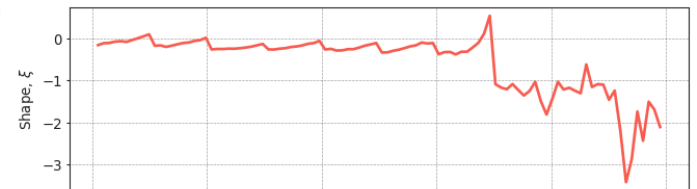
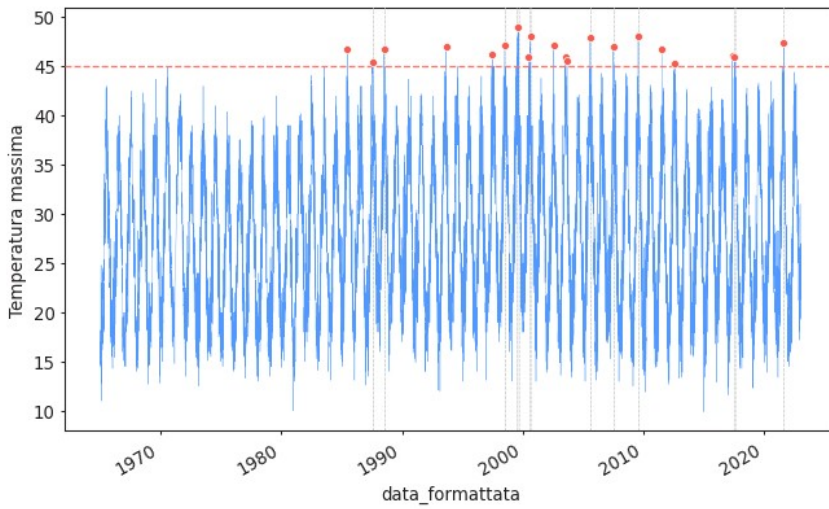


SUD ITALIA:



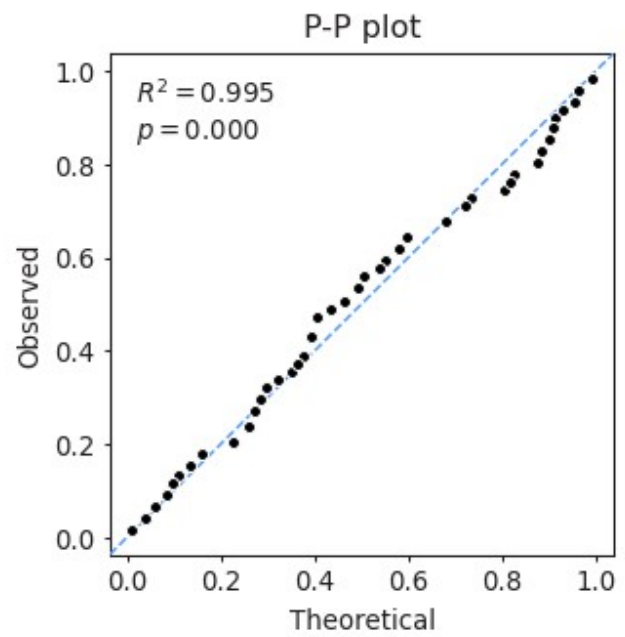
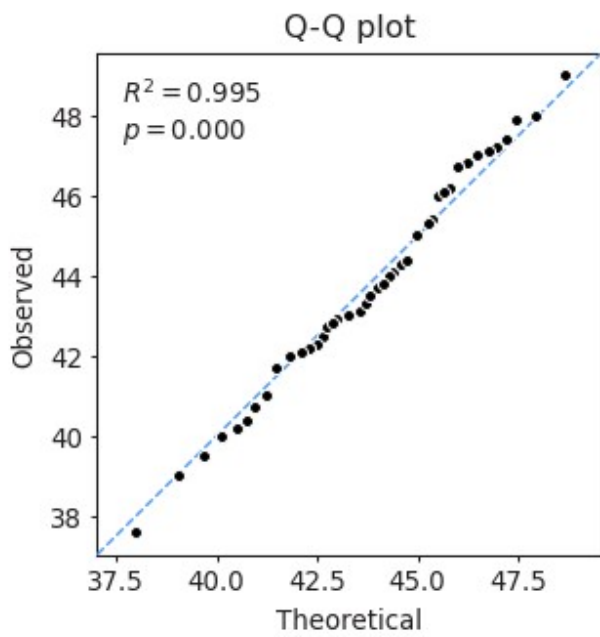
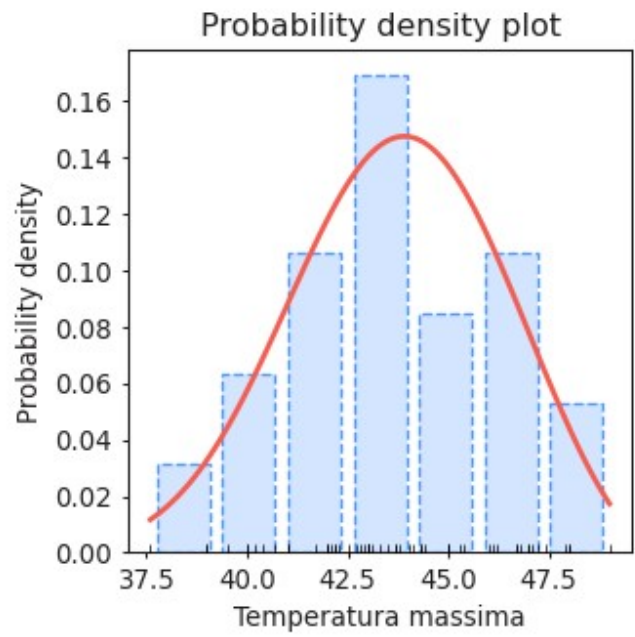
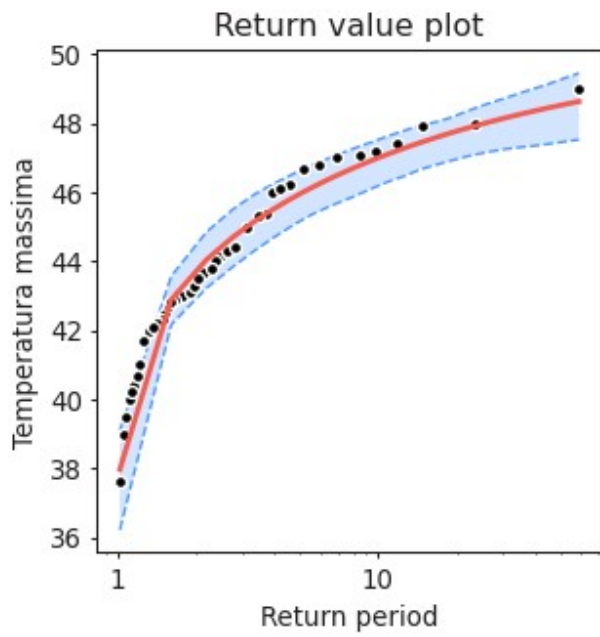
Le 10 temperature più alte registrate nel Sud Italia per ogni anno dal 1965 al 2022:

1999-08-16	49.0
2009-07-23	48.0
2000-08-18	48.0
2005-07-31	47.9
2021-08-11	47.4
1998-07-03	47.2
2002-07-24	47.1
1993-08-27	47.0
2007-06-25	47.0
1985-06-16	46.8



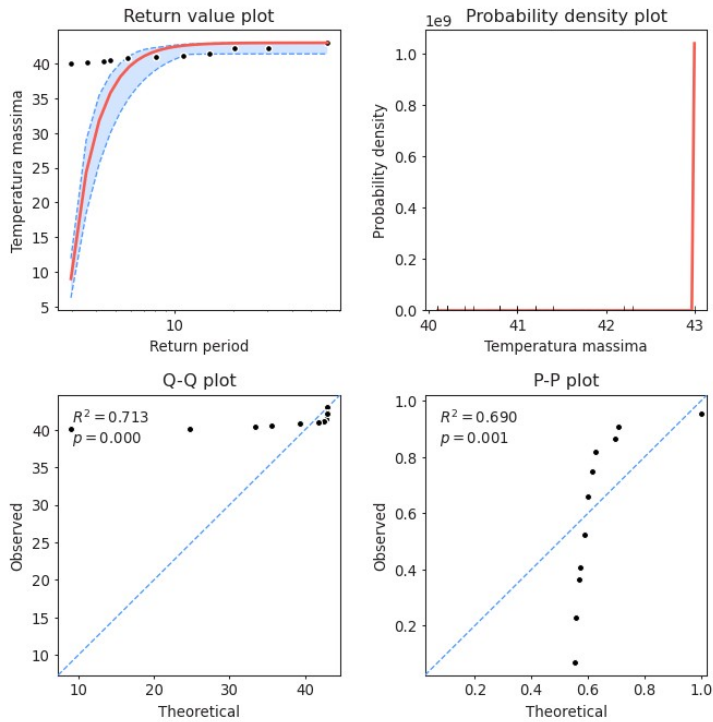
return period	return value	lower ci	upper ci
3.0	44.900603	44.048783	45.747107
6.0	46.259847	45.388434	46.950156
10.0	46.994771	46.150224	47.599906
15.0	47.472927	46.646720	48.004604

data	Temperatura massima	exceedance probability	return period
1999-08-16	49.0	0.016949	59.00
2009-07-23	48.0	0.042373	23.60
2000-08-18	48.0	0.042373	23.60
2005-07-31	47.9	0.067797	14.75
2021-08-11	47.4	0.084746	11.80

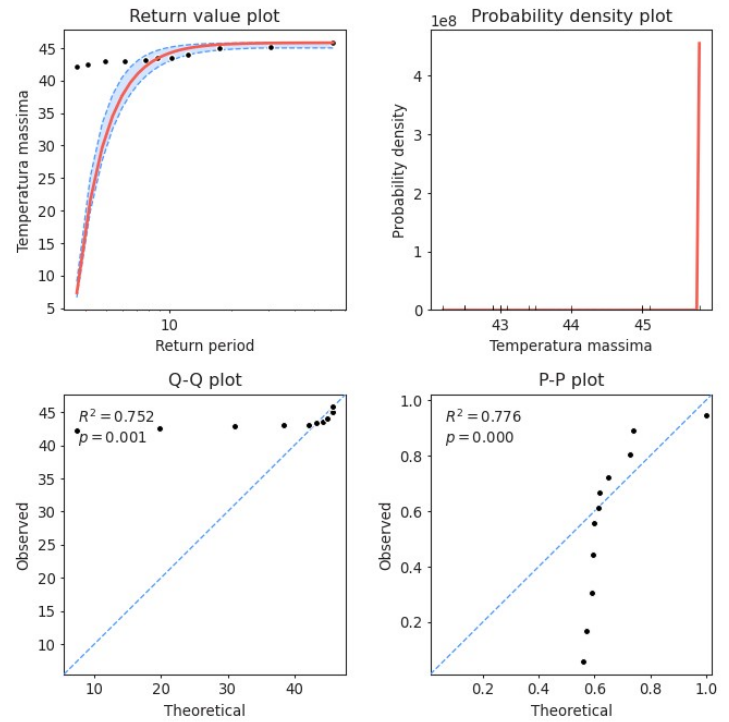


DISTRIBUZIONE GDP:

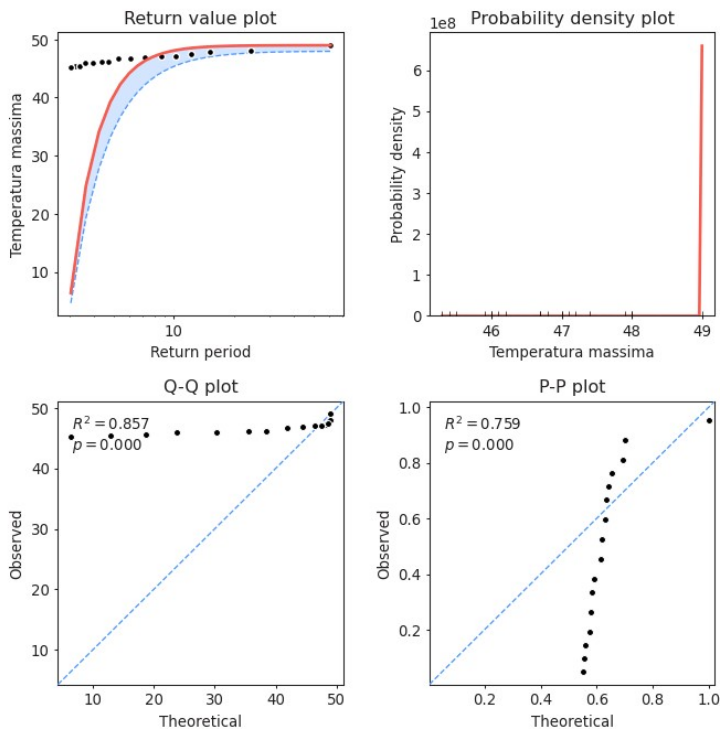
NORD:



CENTRO:



SUD:



Possiamo visivamente notare come il plot dei valori predetti si discosti di molto dai valori effettivamente osservati. Anche il valore del coefficiente di determinazione R^2 (che misura il legame tra la variabilità dei dati e la correttezza del modello statistico) ci indica che i risultati trovati con l'analisi GDP non sono attendibili.