

國立臺灣大學社會科學院國家發展研究所



碩士論文

Graduate Institute of National Development

College of Social Sciences

National Taiwan University

Master Thesis

浮誇新聞與高產榮景：以文字探勘分析大躍進的饑荒

成因

Analysis of The Causes of Famine in The Great Leap

Forward by Text Mining

陳鈺勳

Yu-Hsun Chen

指導教授：鄧志松 博士

Advisor: Chih-Sung Teng, Ph.D.

中華民國 111 年 9 月

September 2022



國立臺灣大學碩士學位論文
口試委員會審定書

本論文係陳鈺勳(R09341026)在國立臺灣大學國家發展研究所完成之碩士學位論文，於民國111年9月15日承下列考試委員審查通過及口試及格，特此證明

口試委員：

鄭志松

(指導教授)

謝吉隆

周嘉良

所 長：

劉輝煌

謝辭

隨著論文的完成，短短兩年的碩士生涯也將畫下句點。儘管求學之路非常艱辛，但同時也收穫滿滿。特別感謝我的指導教授鄧志松老師，除了課堂上學理知識的教學外，也給了我擔任研究助理的機會，讓我從旁學習如何進行研究，更在我撰寫論文遭遇瓶頸時提供建議與指正，正因有老師的幫助，這篇論文才能夠順利完成。也非常感謝在撰寫論文階段曾經幫助過我的課堂老師、同學和研究室的夥伴們，沒有你們我絕對無法在兩年內準時完成學位論文。最後我要感謝我的家人與女友，當我因為課業壓力而心情低落時，你們總是在背後默默的給我支持與鼓勵。我由衷地感謝你們對我的付出，此篇論文我將獻給這一路上曾經幫助我、支持我的所有人，謝謝你們。

摘要



本研究主要探討「大躍進」時期，不同地區的浮誇新聞發佈數量對該地區饑荒死亡人數之影響。從過去文獻可以得知，「大躍進」時期中央常下達遙不可及的高產量任務，導致地方幹部浮報產量的行為層出不窮。當時中國最大的官方媒體《人民日報》更大肆地對這種浮誇資訊加以宣傳，導致全國吹起了一陣「浮誇風」，社會充滿著大量虛假的浮誇新聞。浮誇訊息的傳播不僅讓中央領導者誤以為當時「大躍進」政策正往好的方向發展，進而逐步提高各地糧食生產與徵購的目標，而幹部為了完成因「浮誇風」不斷提高的糧食生產任務，開始不擇手段地從農民身上榨取糧食，中央更因有些地區無法達成生產目標而懷疑當地農民「瞞產私分」才導致缺糧，於是引發一系列官員暴力掠奪糧食的整風，最終各地開始爆發饑荒，大量人民的死亡反映了「浮誇風」的後果。

本研究從《人民日報》官方資料庫擷取了 1958 年至 1962 年發佈的新聞報導，經過資料清理與文本相似性計算篩選出浮誇新聞文本。使用關鍵字分析這些浮誇新聞後發現，《人民日報》所報導的糧食浮誇新聞主要是利用「豐收」、「高產」、「破紀錄」等誇大的形容詞，以及「萬畝」、「萬斤」、「千萬噸」等巨量的單位來塑造當時社會繁華的盛況，其最終目的是為了激勵各地幹部的生產進度以及強化民眾對中央政策的認同感，使「大躍進」政策可以順利執行。

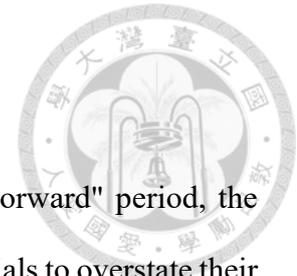
得知浮誇新聞之定義與組成後，本研究透過字典搜索方法將大量浮誇新聞以省級行政區為單位進行分類，得出不同省份的浮誇新聞數量，藉以衡量當地「浮誇風」的氾濫程度。以「浮誇風的氾濫使中央誤信高產狀況而提高生產目標，導致一系列打壓農民、榨取糧食的行為發生，最終饑荒死亡人數增加」做為影響途徑，假設「浮誇新聞越多會導致饑荒死亡人數越多」。為了驗證該假設，本研究將以「大躍進」五年期間各省的「浮誇新聞發佈百分比」作為解釋變數，各省份「人均饑荒非正常死亡百分比」作為被解釋變數進行迴歸分析。研究發現「不同

地區的浮誇新聞數量對饑荒非正常死亡人數有顯著的正向影響」然而有少數饑荒特別嚴重的地區，如四川、安徽則會因為其死亡人數過多而被模型嚴重低估，這表示模型較無法解釋饑荒特別嚴重的地區。



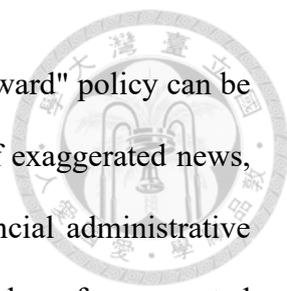
關鍵字：大躍進、大饑荒、文字探勘、中共官媒、迴歸分析

Abstract



It can be seen from past study that during the "Great Leap Forward" period, the central government often require high-yield tasks, causing local officials to overstate their output. At that time, the People's Daily, China's largest official media, promoted this kind of exaggerated information more aggressively, causing a "wind of exaggeration" across the country, and society was full of a lot of fake exaggerated news. The exaggerated information not only made the central government believe that the "Great Leap Forward" policy was going well at that time, and gradually increased the goals of grain production and procurement in various regions, but also the officials were in order to complete their tasks, began to snatch food from farmers by irregular means. The central government even suspected that local farmers were "hiding food" because some areas could not achieve production goals, which led to the shortage of food. Then led to a series of rectification in which officials violently looted food, and eventually famines began to occur in various places, and the deaths of large numbers of people.

This study finds reports published from 1958 to 1962 from the official database of People's Daily, using data cleaning and text similarity screening, and finally classified food-related exaggerated news in different regions by provincial administrative regions. First of all, in order to understanding of the composition and structure of exaggerated news, we can see from the results of keyword analysis that the exaggerated grain news reported by People's Daily mainly uses exaggerated adjectives such as "bumper harvest", "high yield" and "record breaking", and Exaggerated unit quantifiers to shape the prosperity of the society at that time. Through keyword analysis, we can understand the context and conceptual structure of exaggerated news. The purpose is to promote the production progress in various places, communist thought, and to strengthen the public's



sense of identity with the central policy, so that the "Great Leap Forward" policy can be implemented. After understanding the definition and composition of exaggerated news, this study classified a large number of exaggerated news by provincial administrative regions through the dictionary search method, and obtained the number of exaggerated news in different provinces, so as to measure the prevalence of local "exaggerated wind". Using exaggerated news to convince the central government to believe in high production conditions and increase production goals, a series of behaviors that hurt farmers, snatch food, and eventually increase the number of deaths from famine. To test this hypothesis, this study will use the percentage of exaggerated news releases in each province during the five-year period of the Great Leap Forward as the independent variable, and the per capita famine death rate per capita in each province as the dependent variable for regression analysis. The study found that the number of exaggerated news in different regions has a significant positive impact on famine deaths. However, there are a few regions with particularly severe famine, such as Sichuan and Anhui, which are seriously underestimated by the model because of their excessive number of deaths, which means that the model is less able to explain areas with severe famine.

Keywords: Great Leap Forward, Three Years of Great Famine, Text Mining, CCP Official Media, Regression.

目 錄



論文口試委員審定書	i
謝辭	ii
摘要	iii
Abstract	v
目 錄	vii
圖表目錄	ix
第一章 緒 論	1
第一節 研究緣起	1
第二節 問題意識與主要研究問題	2
第二章 相關理論回顧與文獻探討	5
第一節 「大躍進」與「大饑荒」相關研究	5
壹、「大躍進」的發起與影響	5
貳、三年大饑荒的成因與死亡人數研究	7
第二節 中國大陸媒體之研究	12
壹、黨與媒體的關係	12
貳、「大躍進」時期媒體之傳播途徑	13
第三節 浮誇風與饑荒的關聯	15
第四節 文字探勘之技術與應用	17
第三章 研究設計	20
第一節 研究方法與架構	20
壹、資料前處理	20
貳、糧產浮誇新聞的篩選	21
參、浮誇新聞的關鍵字分析	27
肆、各省的浮誇新聞數量	28
伍、迴歸分析	29
第二節 研究對象與資料來源	30
壹、各分析方法之研究對象	30
貳、資料來源	31
第三節 分析工具	31
第四節 論文結構與章節安排說明	32
第四章 研究結果	34
第一節 浮誇新聞的篩選與分類結果	34
壹、資料前處理	34
貳、浮誇新聞之表述方法	35
參、新聞地區別分類方法之定義與結果	38



第二節 浮誇新聞的關鍵字分析	39
壹、關鍵字提取.....	39
貳、關鍵字頻分析.....	40
第三節 迴歸分析	43
壹、樣本敘述統計.....	43
貳、分析結果.....	46
第五章 結論.....	50
第一節 主要研究發現.....	50
壹、浮誇新聞的表述方式.....	50
貳、浮誇新聞與饑荒嚴重程度之關係	51
第二節 後續研究建議.....	51
參考文獻.....	53
壹、中文.....	53
貳、英文.....	55

圖表目錄



表 1 1957-1965 年中國官方公布戶口登記人口	8
表 2 國外學者推估之歷年死亡率	8
表 3 國外學者推估的非正常死亡人數 (單位：萬人)	9
表 4 楊繼繩估計的歷年非正常死亡人口 (單位：萬人)	10
表 5 曹樹基的各省災害統計	11
表 6 模型參數設定與篩選結果比較	26
表 7 第一類浮誇新聞簡要範例	35
表 8 第二類浮誇新聞簡要範例	36
表 9 「大躍進」五年期間各省浮誇新聞數量統計結果	38
表 10 關鍵字出現頻率前 20 名 (單位：次)	40
表 11 樣本敘述統計表	43
表 12 分析結果	47
圖 1 媒體傳播途徑流程圖	15
圖 2 CBOW 與 Skip-gram 架構	22
圖 3 PV-DM 架構	23
圖 4 Transformer 架構	24
圖 5 Multi-Head Attention 架構	25
圖 6 文本相似性計算流程	27
圖 7 地區別浮誇新聞分類流程	28
圖 8 研究流程與架構	30
圖 9 歷年浮誇新聞數折線圖	39
圖 10 關鍵字聲量折線圖	41
圖 11 關鍵字文字雲圖	42
圖 12 「浮誇新聞數百分比」與「人均非正常死亡百分比」盒鬚圖 ..	44
圖 13 人均非正常死亡數百分比長條圖	45
圖 14 浮誇新聞數百分比長條圖	46
圖 15 迴歸線&散佈圖	47



第一章 緒論

第一節 研究緣起

改革開放以前，中國官方媒體是傳遞中央政策的關鍵中介，以黨的絕對領導為傳播核心，具有高度的社會影響力。做為官方的宣傳工具必須以黨國為主體，將所期望引導之內容進行傳播，以達到國民接受主體思想，產生正確輿論之目的。任何新聞報導皆須依附黨中央的政治目的，將輿論導向政策的基本規範。「一黨專政」的政權特性，就是最大力量控制媒體和輿論，藉以引導報導內容，操控社會民意以維護政權穩定（陳福成，2000）。

自中共建黨以來，常常以「官媒」作為特定的訊息傳播管道推動政策，並利用媒體影響社會大眾對公共事務的關注，發揮意識型態的導向作用與維持民意發展和社會正向的新聞活動（王毓莉，1998），最終達成黨所期盼的戰略目標。加強媒體對輿論的導引或以行政命令管制境內新聞，是中共面對負面事件或者重大問題一貫的手法（杜聖聰，2008）。過去中共將傳播媒體定位為集體的鼓動者、宣傳者及建設共產主義的教育者，擁有強烈的工具性，是一種意識形態機構（王毓莉，2009）。儘管在改革開放後，媒體的自主性有所增加，但許多議題仍會受到黨的控制與限制。媒體至今仍常常因為黨的政治因素而失去報導應有的客觀、理性與科學的態度。違反這些要素的不實報導不僅深刻影響社會輿論，亦會影響政府的施政方針，因此研究極端媒體行為所導致的後續影響就顯得十分重要。

1958年中國共產黨所推行的「大躍進」運動就是一個極端的案例。在20世紀發生的「大躍進」運動中，中共中央官媒《人民日報》扮演著「引風吹火」的角色（黃永鋒、李代友、韓華淑，2015）。這項探索社會主義建設道路的嶄新嘗試最終釀成了1959年至1961年的「三年大饑荒」。這場被廣泛認為是人類歷史上最規模的饑荒，在三年間非正常死亡人數被預估高達1500萬至5500萬人（楊繼



繩，2009)。這場災難不僅是錯誤政策所導致的後果，受政治因素影響而報導片面、失真、浮誇的新聞媒體更是其中的幕後推手之一。《人民日報》作為當時最具影響力之官方傳媒機構，其發表的報導足以直接影響每一場政治運動的方向。通過《人民日報》所發佈的報導、宣傳與口號，常會形成社會輿論。這些報導的功能除了向國民傳達當前中央的政策方向之外，也能反饋社會當前的政策執行成果供中央高層決策者參考，甚至以該報導作為制定未來施政方針之依據。例如，在1958年的北戴河會議中，毛澤東相信了《人民日報》發佈的生產高指標和浮誇報導，誤以為當前政策正往好的方向發展，於是加強了「工業大躍進」、「農業大躍進」與「人民公社」的施政強度與規模，造成饑荒的延長與國內經濟大幅衰退。1961年「大躍進」結束後，時任中華人民共和國主席劉少奇對「大躍進」的教訓進行了反思。他多次批評《人民日報》的行為，並在《關於人民日報工作的談話》中提及：「大躍進的責任，中央領導負一半，人民日報領導負一半」，並指出：「有報紙的害處比沒報紙的害處更大。錯誤的宣傳造成很大惡果，人民日報宣傳生產建設的浮誇風，瞎指揮風的經驗和片面性理論，影響很壞」。這些言論間接證明了當時浮誇不實的報導有著巨大影響力，且是造成數千萬人喪生的重要原因。有鑒於此，本研究將以大躍進時期《人民日報》所發表之報導著手，以文字探勘的方式對報導進行分析，透過量化該時期所發佈的新聞報導，全面性的探討當時浮誇的報導是如何影響饑荒。

第二節 問題意識與主要研究問題

1957年，蘇聯發射了第一顆人造衛星使中國領導們為之振奮且憧憬，往後便把社會主義建設中非常傑出的成就稱之為「放衛星」。1958年的「大躍進」專注於生產的高速度與高產量，當時地方幹部常為了自保或升遷而上報與實際不符的生產成果。《人民日報》甚至將這些浮誇的成果統一公布在《比一比》專欄中比較各省的生產進度，這種行為使得各省幹部因為壓力而爭先恐後大搞浮誇。虛報



產量的訊息透過媒體報導後形成了浮誇的「高產衛星」新聞，導致當時的報紙不斷有糧食高產「放衛星」的喜訊。

過去探討「大躍進」災難責任的相關研究皆指出，除了政策方面的錯誤之外，中央的高估產量以及反右傾思想；基層幹部千方百計地迎合並虛報產量；媒體不揭露真相甚至大放衛星，皆是引發大饑荒的主要原因。另外，當時的媒體並不只從事傳達中央思想與政策的單純工作，更常因為政治壓力或利益而聽信地方幹部所上報的高產量成果，進而散播「繁榮社會」與「政策推行順利」的不實報導，藉此營造社會主義建設十分成功的氛圍（王官德等，2008）。此時的社會充斥著高產的一片榮景，以致中央也聽信了這些虛報成果，逐步提高糧食徵購量和生產指標，加速「大躍進」施行進度。高產神話帶來兩種直接後果，分別為毛澤東誤信這片榮景，擔心糧食吃不完而鼓勵民眾大吃大喝；國家的徵購指標隨著浮誇虛報的產量而水漲船高。而現實是各地區的糧食嚴重不足，徵購任務難以完成，地方幹部只好把徵購糧食和反右傾思想結合在一起，用反右傾的政治壓力最大化的從農民身上榨取足夠的糧食。當時有些地方已經出現饑荒，然而中央卻因為浮誇風的影響，誤以為各地存在「瞞產私分」才導致缺糧，進而引發一系列暴力掠奪糧食的整風，最終各地饑荒越演越烈，農民以生命為代價承擔了浮誇風的後果（楊繼繩，2009）。

綜合以上所述，各地區的幹部可能因為徵購任務的壓力而透過虛報產量來迎合中央，不同地區浮誇風氣的氾濫程度則會因幹部的政治狂熱或對中央忠誠度而不同。這些虛假訊息透過《人民日報》發佈成新聞並演變成全國性的浮誇風。其更創造了各地的生產榮景並誤導中央認知，導致一系列打壓農民、榨取糧食的行為發生，最後爆發大饑荒。因此本研究假設「地區的浮誇新聞越多會導致饑荒死亡人數越多」。為了驗證這項假，本研究以浮誇新聞作為切入點探討「大躍進」運動下不同地區媒體的浮誇程度與饑荒嚴重程度之關聯，利用文字探勘（Text Mining）技術來量化「大躍進」期間的報導，計算出各地區浮誇新聞的數量作為

衡量各地「浮誇風」程度的依據。饑荒的嚴重程度本研究將以過往學者預估的「饑荒非正常死亡人數」作為衡量。

為全面客觀的探討研究問題，本研究首先會針對時間範圍內《人民日報》發佈的所有新聞開始著手分析，透過程式有效地從大量新聞中區分出浮誇新聞，再進一步將大量的浮誇新聞進行地區（省）的分類，計算該地區每年的浮誇報導共有幾篇，並以迴歸分析來估計「浮誇新聞」對饑荒非正常死亡人數的影響效果，最後觀察迴歸線散佈圖和比對資料的分佈與趨勢得知浮誇報導越多的地區是否饑荒會越嚴重。

研究時間範圍方面，由於「大躍進」結束後一直到1961年初，中共中央才開始正視其所造成的錯誤，並於1962年召開的《七千人大會》，總結這項政策執行以來的經驗教訓與提出修正錯誤之方針。因此本研究假設「大躍進」的影響直到1962年才真正告一段落，是故研究的時間範圍將設定為1958年1月至1962年12月。透過分析這五年的報導內容、參考當時的歷史重要事件與過去學者的觀點，梳理出本研究之問題核心。

第二章 相關理論回顧與文獻探討

第一節 「大躍進」與「大饑荒」相關研究



壹、「大躍進」的發起與影響

中共建政七年後，基本上已將中國改造成一個社會主義國家，然而當時的中國相較於歐美列強仍顯得貧窮且落後（王官德等，2003），儘管第一個五年計畫已經為中國建立工業化技術的初步基礎，但現代化工業仍舊停滯不前，農業集體化也沒能形成技術革命（Kirby, 2011）。

1956年4月25日，毛澤東提出〈論十大關係〉的報告，論述了社會主義革命和社會主義建設的關係。梁柱（2013）指出，〈論十大關係〉的中心思想就是從中共自己的國情出發，走自己的路，提出適合中共國情的「多快好省地建設社會主義總路線」的基本思想，是探索中國社會主義建設道路的第一步。1956年5月2日，毛澤東在最高國務會議第七次會議上正式提出實行「雙百」運動。在這場運動中，毛澤東透過肅清大量知識分子達到殺雞儆猴的效果，最終使多數右傾或感於發表的知識分子噤若寒蟬，重新獲得黨的控制權（王官德等，2003）。而中共中央於1957年10月初，批准毛的激進農業發展規劃，並準備嘗試更進一步的社會主義化行動（閻茂旭、劉璇，2006：50-52）。「雙百」運動期間，中共同時拋棄了蘇聯經濟的發展模式，開始走中國式的社會主義建設路線，與此同時，中共也試圖脫離史達林主義的影響（陳永發，1998）。1956年9月15日至27日，中國共產黨召開第八次全國代表大會。毛澤東指出此次大會的主題是：「總結前七次大會以來的經驗，團結全黨，團結國內外一切可能團結的力量，為建設一個偉大的社會主義國家而奮鬥」（中共中央文獻研究室，2009），此次會議中的發言皆圍繞在探索適合中國實際情況的社會主義建設路，使中共探索社會主義向前邁進一大步。



1957年11月2日，毛澤東率領中國代表團第二次訪問蘇聯。此次出訪，毛澤東受到蘇聯成功發射人造衛星的消息所震撼，故認為蘇聯在科技方面後來居上美國，此事件更加深毛澤東對於社會主義國家優於資本主義國家的觀念（陳永發，1998）。此外在訪蘇期間，蘇聯總書記宣布15年內蘇聯要超越美國，毛澤東也因此受到刺激，並隨後宣告「中國要在十五年左右，在鋼鐵等主要工業產品產量方面趕上或超過英國」。同年12月2日，劉少奇代表黨中央，公開宣布十五年內趕超英國的口號，以致後續農村掀起了一系列以興修水利、養豬積肥和改良土壤為中心的農業改革，正式揭開了「大躍進」的序幕。

Garnaut (2013)指出，1958年至1960年的「大躍進」是中共探索社會主義建設道路的嶄新嘗試，這項運動有其歷史發展背景，首先是反右傾鬥爭的影響，毛澤東認為鬥爭的勝利使得中國更接近共產主義的天堂。另外，整風運動中有些工廠、農村出現了大量產出的現象，因此更加堅信1956年的“反冒進反錯了”。1958年5月5日，八大第二次會議在北京舉行。大會正式通過毛澤東的倡議“鼓足干劲、力爭上游、多快好省地建設社會主義”的總路線，肯定了全國出現的「大躍進」型式，並將鋼產量計畫指標由1200萬噸調高至3000萬噸。毛澤東在會上講話，強調要破除迷信，解放思想，發揚敢想敢說敢做的創造精神。此次會議後，全國再次掀起「大躍進」的第二波高潮。（楊繼繩，2009）。1958年後，毛澤東因為受到媒體的浮誇報導所蒙蔽，誤信「大躍進」正在往好的方向發展（茹倩，2013），故為了加速實現全面社會主義與超越列強的目標進度，於1958年5月開始大辦「人民公社」。

1958年8月17日至30日，毛澤東在北戴河召開中共中央政治局擴大會議（簡稱北戴河會議），正式定義「大躍進」的各項主要計畫。會中討論了未來的國民經濟計畫、第二個五年計畫以及當前的工農商運作問題，同時也公布了《中共中央政治局擴大會議號召全黨全民為生產1070萬噸鋼而奮鬥》的公報以及《中共中央關於在農村建立人民公社問題的決議》（陳永發，1998）。毛澤東相信生產



的高指標和浮誇消息，故要求繼續擴大「人民公社」，並預言大約再三至六年就可以將中國打造成全民所有制的共產主義社會。直到 1958 年 11 月，全國已建成大約 26,500 個公社，佔所有農村家庭的 99.1%，每個公設平均容納 4,756 戶。一直到 1959 年底，將近 4 億人口被迫在 391,900 個公社食堂吃飯，農民無法搬出公社，且天天在工作分配的公平性問題上與幹部發生衝突（楊繼繩，2009）。儘管「人民公社」與「大躍進」都存在問題，但毛澤東仍主觀認定合作社的規模越大、公有化程度越高，就越能促進生產，且效果非常卓越，是社會主義過渡的基本單位（董芹，2007），故決定把工、農、商、學、兵組成為一個大公社，為指導農民加速社會主義建設，提前建成社會主義的基本方針。

貳、三年大饑荒的成因與死亡人數研究

1959 至 1961 年間，中國發生了一場全國性饑荒，這場饑荒被廣泛視為人類歷史上規模最大的饑荒，也被認為是人類歷史上最嚴重的人為災難之一，有專家學者認為饑荒不僅限於這三年（Dikötter, 2010），而是延長到 1962 年。中華人民共和國將這場災難命名為「三年困難時期」，而國外學者則常稱其為「三年大饑荒」、「中國大饑荒」或「大躍進饑荒」。其發生成因仍然是學術界一個激烈爭論的話題。有些專家將其歸咎於自然災害和極端惡劣的天氣；而另一些專家則聲稱這是「大躍進」政策錯誤所導致的結果；另有一派學者認為，「三年大饑荒」主要是毛澤東所導致，通常這一派學者將毛澤東的角色描繪成：邪惡、喜歡殺人、無情的政治家等形象。儘管毛澤東知道推行「大躍進」可能會造成大量死亡，但仍為了建設社會主義國家的目標而不顧一切（Karlsson & Schoenhals, 2008）。針對這幾種說法，Houser 等人（2009）對大饑荒的成因進行了全面性的探討，其研究表明，饑荒期間某些地區確實有出現過極端氣候，且這些地區的氣候確實對饑荒的發生有顯著的影響。然而，以全國來說饑荒的成因仍然只有一小部分原因能歸咎於天氣，因為受害者人數太多，不可能僅由天氣條件所造成。饑荒期間的死亡



人數，中共官方與各方學者皆有不同看法，中國大陸官方曾於 1953 年和 1964 年進行了兩次全國人口普查，並於國務院批准後首次於 1984 年出版的《中國統計年鑒》(中國國家統計局，1984)，對外公布了 1949 年到 1982 年戶口登記的年度人口數字(如表 1)。

表 1 1957-1965 年中國官方公布戶口登記人口

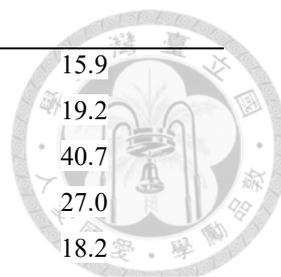
年份	年末總人口 (萬人)	出生率(‰)	死亡率(‰)	自然增加率
1957	64,653	34.03	10.80	23.23
1958	65,994	29.22	11.98	17.24
1959	67,207	24.78	14.59	10.19
1960	66,207	20.86	25.43	-4.57
1961	65,859	18.02	14.24	3.78
1962	67,295	37.01	10.02	26.99
1963	69,172	43.37	10.04	33.33
1964	70,499	39.14	11.50	27.64
1965	72,538	37.88	9.50	28.38

資料來源：(中國國家統計局，1984)。

Frank (2010) 根據歷史學家與統計學家之研究，大膽的預估這三年的饑荒非正常死亡人數超過 4500 萬人。美國人口學家 Banister (1991) 曾使用 1984 年中國公布的千分之一生育率調查數據，重新對出生率與實際死亡率進行估計。其研究為三年大饑荒期間的人口變動提供了重要的基礎。美國人口學委員會主席 Coale (1984) 同樣使用中國公布的千分之一生育率抽樣調查資料和三次人口普查資料重估歷年出生與死亡率，與 Bannister 不同的是，其研究還使用推算出來的歷年出生人數與死亡率(如表 2)，推估出非正常死亡人數。法國國立人口研究所所長 Calot (1985) 也使用相同的數據與類似的方法估算當時中國人口的出生率和非正常死亡人口，這三位學者所估計的非正常死亡人數如表 3 所示。

表 2 國外學者推估之歷年死亡率

年份	Bannister (%)	Coale (%)	Calot (%)
----	---------------	-----------	-----------



1958	20.65	20.4	15.9
1959	22.06	23.3	19.2
1960	44.6	38.8	40.7
1961	23.01	20.5	27.0
1962	14.02	13.7	18.2

資料來源：本研究參照 Hong et al, 2013:20，自行整理。

表 3 國外學者推估的非正常死亡人數（單位：萬人）

年份	Bannister	Coale	Calot
1958	300.1	271.0	130.3
1959	396.8	463.2	214.0
1960	1833.5	1468.9	1632.8
1961	456.7	277.6	721.4
1962	-	-	152.4
合計	2987.1	2481	2850.9

資料來源：(楊繼繩，2009：457)。

- 說明：1. 非正常死亡人數為正常死亡率減當年實際死亡率再乘以該年總人口。
2. 三位學者對饑荒持續時間有不同見解，Calot認為饑荒持續至1962年。

中國學者楊繼繩（2009）透過其新華社資深記者的身分，調閱了各省的官方機密資料，並以 1955、1956、1957、1963、1965、1966 等年度死亡率之算術平均推論出「大躍進」期間的正常死亡率，最後與饑荒各年之實際死亡率相減，即為當年之非正常死亡率，如此即可使用其調閱之數據來推斷出大饑荒期間各省的非正常死亡人數（如表 4），並得出 1959 至 1962 年共有 2,098 萬人為非正常死亡之結論。而地理學者曹樹基（2005）採用人口學和地理學的方法，以清朝虛擬的「府級」行政區為單位，通過各地的地方誌所記載的資料來計算出大饑荒期間災前、災後的人口增長速度，確定各地災前災後人口後即可求得非正常死亡人口，最後考慮該省的人口、縣府數量，依照饑荒之災害嚴重程度做出排序，並指出三年大饑荒期間，全國非正常死亡人口合計約 3,246 萬（如表 5）。蔣正華和李南

(1998) 則以當時國民的生命表（也稱死亡率精算表）為基礎，估算出歷年的死亡人數，並得出三年饑荒累積的非正常死亡人數約 1,700 萬人。可以明顯看出其估計的非正常死亡人數相較其他學者低了不少，儘管蔣正華的 1,700 萬為國內的主流所接受，並被視為權威數據。但學術界認為這個數字嚴重低估，且計算方式有諸多部分受到質疑，因此普遍不被採用（楊繼繩，2009）。

表 4 楊繼繩估計的歷年非正常死亡人口（單位：萬人）

地區	1958年	1959年	1960年	1961年	1962年	總計
四川	98.53	253.37	299.42	123.04	22.65	797.01
安徽	9.46	24.12	192.7	-	-	226.28
山東	14.3	43.5	71.18	43.75	11.71	184.43
河南	9.89	17.18	139.38	-	-	166.45
湖南	-	4.93	64.67	20.72	-	90.32
貴州	1.19	10.05	63.96	14.34	-	89.59
廣西	1.43	14.07	40.23	18.23	-	73.96
甘肅	14.45	9.91	40.1	2.21	-	66.67
江蘇	-	17.52	33.97	12.3	-	63.79
遼寧	2.42	9.89	9.36	24.63	1.77	48.07
廣東	1.8	11.49	24.38	7.72	2.63	48.02
雲南	15.12	8.28	24.04	-	-	47.44
湖北	-	13.11	34.28	-	-	47.39
河北	1.78	6.96	20.44	12.08	-	41.26
福建	-	-	11.56	6.26	0.58	18.4
江西	-	3.32	9.49	0.48	-	13.29
吉林	0.09	5.67	1.45	4.17	1.4	12.78
浙江	-	2.61	6.73	1.43	-	10.77
青海	0.61	1.56	7.77	0.35	-	10.29
黑龍江	-	5.71	2.13	3.29	-	11.13
山西	0.24	2.06	4.46	1.11	-	7.87
陝西	-	2.78	2.1	-	-	4.88

寧夏	0.91	1.15	0.79	0.12	-	2.98
新疆	-	3.45	1.62	-	-	5.07
內蒙	-	1.2	-	-	-	1.2
北京	-	-	0.98	1.55	-	2.53
天津	-	0.97	1.29	1.02	-	3.26
上海	-	-	0.49	1.44	1.02	2.95
總計	172	475	1109	300	42	2098

資料來源：(楊繼繩，2009：451)。

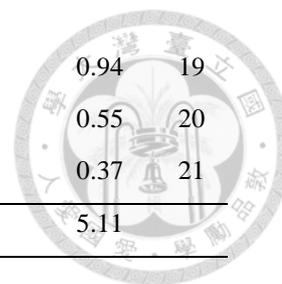
說明：1.楊繼繩以官方資料為基礎，計算之非正常死亡人數。

2.無值欄位代表當年度之死亡人數為合理範圍。

表 5 曹樹基的各省災害統計

省名	全部府 數	災區府 數	全部縣 數	災區縣 數	災前人口 (萬)	非正常死亡人 口(萬)	比例 (%)	排 序
安徽	13.0	13.0	74.6	73.6	3446.0	633.0	18.37	1
四川	26.0	23.0	201.0	188.0	7191.5	940.2	13.07	2
貴州	16.0	15.0	81.0	74.0	1706.4	174.6	10.23	3
湖南	17.0	17.0	96.0	91.0	3651.9	248.6	6.81	4
甘肅	13.0	12.0	108.0	87.3	1585.7	102.3	6.45	5
河南	13.0	13.0	128.7	106.7	4805.8	293.9	6.12	6
廣西	12.0	12.0	75.0	72.7	2009.9	93.1	4.63	7
雲南	22.0	22.0	132.0	112.0	1918.3	80.4	4.19	8
山東	12.0	12.0	140.0	126.0	5343.5	180.6	3.38	9
江蘇	12.0	11.0	84.0	63.0	5296.3	152.7	2.88	10
湖北	11.0	11.0	79.8	72.6	3074.9	67.5	2.20	11
福建	11.0	11.0	70.0	56.0	1547.9	31.3	2.02	12
遼寧	-	-	-	-	1931.0	33.0	1.71	13
廣東	15.0	12.0	110.0	72.0	3839.3	65.7	1.71	14
黑龍 江	-	-	-	-	1564.0	19.0	1.21	15
河北	18.0	17.0	183.0	135.0	5544.6	61.0	1.10	16
江西	14.0	13.0	89.0	51.0	1713.6	18.1	1.06	17
陝西	12.0	4.0	101.0	30.3	1832.0	18.7	1.02	18

吉林	-	-	-	-	1280.9	12.0	0.94	19
浙江	11.0	10.0	91.0	51.4	2570.3	14.1	0.55	20
山西	-	-	-	-	1621.1	6.0	0.37	21
合計	248.0	228.0	1844.1	1462.6	63474.9	3245.8	5.11	



資料來源：(曹樹基，2005：附錄 2)

說明：1. 曹樹基以地方誌記載資料為基礎計算之非正常死亡人數。

2. 無值欄位代表當年度之死亡人數為合理範圍。

第二節 中國大陸媒體之研究

壹、黨與媒體的關係

中國媒體常扮演著鼓動者、宣傳者以及共產主義教育者等多種角色（王毓莉，1998）。改革開放後，中國的傳播媒體雖經多次的新聞改革，但媒體的自主性仍然不及歐美等自由民主國家。儘管《中華人民共和國憲法》賦予公民享有言論、出版自由，但「黨管媒體」的說法依然存在（潘忠黨等，2008）。過去研究指出，傳播媒體歷來常被中國共產黨作為政治工具應用，且傳播的信息必須在政治立場上和中央保持一致（韋旭，2008）。其中，本文欲研究之標的《人民日報》，自中共建黨以來一直是規模最大之傳播媒體，影響力甚鉅。

在核心的媒體領域，國家無法接受媒體與文化機構產權的私有化，且禁止國外的私人資本進入國內傳媒的關鍵領域（潘忠黨等，2008）。中共透過社會公眾中的意見領袖或參與事件和表達意識的活躍人群來控制輿論走向。這種輿論具有一定的權威性和導向性，能夠較大範圍的傳播事實、意見和觀點（吳柳林，2011）。韋旭（2008）指出，中國社會存在著主流媒體營造的“媒體輿論”以及群眾議論紛紛的“公眾輿論”。媒體輿論能夠導引公眾輿論，使其更具吸引力和感染力，間接影響群眾的思想和行為。Guan (2018)認為框架理論可以廣泛應用於政治傳播研究，亦可以解釋中共政府利用媒體傳播來構建輿論的行為。框架的概念強調一個話題對不同面向的影響力，它可以賦予一個事件或問題特定的意義，從而

影響觀眾對話題的態度 (Vreese, 2005)，中共政府能通過黨的權力來維護官媒信息傳播的合法性。相較於西方媒體，中國的媒體能夠更有效建構輿論。在毛澤東時代，黨的機關就已經幾乎壟斷政治和社會生活方面的新聞框架 (吳柳林, 2011)。

目前中國有 18 家中央媒體，包括新華社、解放軍日報、光明日報、中央人民廣播電台等。在 18 家中央媒體中，《人民日報》無疑是占據主導地位。主要是因為《人民日報》作為中共中央機關報以及其長遠的歷史和演進背景。1992 年，《人民日報》被聯合國教科文組織評為世界十大報紙之一，為當今世界上級別最高的媒體之一。《人民日報》無疑是親政府的，在承載新聞價值和商業利潤的同時，亦肩負著宣傳的社會責任。其報導關注之政治和社會問題，深深影響著國內受眾 (Yan, 2020)。鑒於《人民日報》是國有媒體，受黨直接監管，儘管民眾可能因其立場而對傳達的信息抱持懷疑態度，但因為其規模最大，因此民眾仍深受其發佈的信息所影響。過往許多文獻都可以發現《人民日報》常常發佈帶有政治目的的報導，且報導的內容可以有效控制民眾的行為。在廣東的 SARS 事件爆發初期中央為了掩蓋疫情，使《人民日報》對此事報導的內容不是語意不明就是欲言又止，從而限制了公眾對真實情況的了解，公眾無法做出正確判斷，造成後期 SARS 病例與流行範圍不斷擴大 (方雪琴, 2004)。2020 年 1 月新型冠狀病毒在中國肆虐期間，官方媒體發佈許多以“也許”、“可能”等用詞含糊不清或帶有錯誤資訊的方式描述疫情的傳播途徑或藥品，進而導致人民誤信某些預防措施或特定藥品能夠預防病毒的情況 (Zhang, 2020)。這些研究皆指出，災難事件來臨時，公眾歷來十分關注媒體，且對於災難的理解經常會受限於媒體所創造的框架，最終按照媒體的指示行動。

貳、「大躍進」時期媒體之傳播途徑

「大躍進」時期，人民誤信或者受到媒體影響的情況更為嚴重。姚桂榮



(2012) 指出，當時的新聞傳播媒介主要有：國家通訊社的新華社、中央黨報《人民日報》、黨刊《紅旗》雜誌，以及中宣部主管的《光明日報》等少數媒體，而參與「大躍進」運動的廣大幹部和民眾就是這些傳播媒體的主要受眾。由於當時電視沒有普及，新聞傳播主要是通過報紙或黨刊進行。《人民日報》作為官方的正式傳播媒體，受眾將會對報導內容有先入為主的觀念同時產生輿論導向的效果。

威權時代下傳播者和受眾在「大躍進」期間相互影響，幹部與民眾不斷的受媒體刺激而促使群眾行為不斷加強。相較於改革開放後，當時的媒體是由毛澤東直接控制，新聞的內容帶有更多的政治色彩（董芹，2007）。也因為媒體在「黨委辦報」下運作，若試圖揭露或公開真實的情況，即有可能被扣上了「右傾」或「反三面紅旗」的帽子，這也導致新聞失真的浮誇風情形非常普遍。亦有文獻指出，「大躍進」期間，媒體行為除了影響群眾之外，甚至反向影響中央高層。高產訊息從地方官員發出，媒體進行轉述後反映錯誤的現況給群眾和中央，最終中央接收媒體傳遞的資訊後進而增加施政強度與徵購目標，形成輿論傳播的循環（如圖 1）。「大躍進」一直持續至 1961 年才宣告停止，這場舉國上下的社會運動沒有讓中國進入一個新的偉大時代，反而以災難性的失敗告終。過去研究「大躍進」的歷史學者皆認為這一系列的計畫太過極端，目標太高，甚至遙不可及，Garnaut (2013) 也認為，「大躍進」是中共建黨以來造成社會資本破壞最為嚴重的時期，不僅經濟大幅度倒退，還造成了大量人口死亡。除此之外，「大躍進」也帶來了許多不利的後果，包括饑荒過後國民的健康產生了長期顯著的負面影響（Almond et al., 2007），人民的信仰遭受侵害，導致政府與法律失去了社會的信任（He & Sun, 2016）。

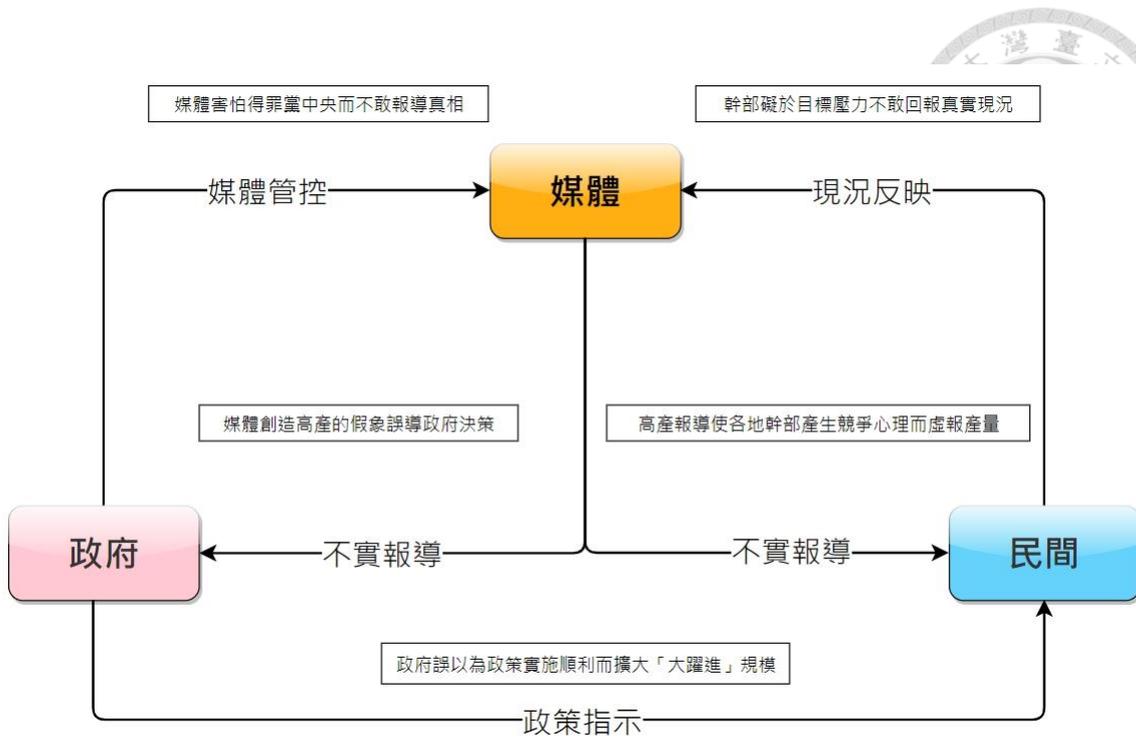


圖 1 媒體傳播途徑流程圖

資料來源：本研究自行整理

第三節 浮誇風與饑荒的關聯

過去文獻指出，中共於 1958 年至 1961 年推動「大躍進」運動是有鑑於西方列強與蘇聯的壓力和刺激。其發展戰略旨在通過高度集中改革，盡速實現工業技術革命 (He, 2012)。然而「大躍進」最終以失敗收場，關於失敗的原因，除了政策上的錯誤，當時的「浮誇」也是造成失敗的主因之一。

浮誇風是指「大躍進」運動中虛報產量以及媒體浮誇不實的風氣。媒體是助長其盛行的最大媒介，浮誇風現象之所以能不斷持續且一發不可收拾，正是因為毛澤東的獨斷專制 (靖鳴、劉銳，2008)。葛玲 (2010) 認為浮誇風背後體現的是「寧左勿右」的政治邏輯，但並非是由下自上生成，而是上級的高產期望導致下層幹部的高產壓力才促成浮誇風的興起。許多研究指出，浮誇風是導致後續大饑荒發生的主要原因之一，由於高產的消息遍布全國，導致各地的徵糧目標不斷



提高，儘管當時多數幹部皆心知肚明這些產量是捏造的，但若跟不上其他地區上報的產量則遭受打壓（鐘萬全，2003）。例如《人民日報》曾在「大躍進」期間發佈《比一比》專欄，在專欄中公布各省的生產紀錄與目標完成進度。使得各省幹部爭先恐後大搞浮誇（姚桂榮，2012）。浮誇風報導氾濫，只報喜，不報憂，使毛澤東聽信這形勢一片大好的報告之後，竟然在全國缺糧的情況下，提出了“糧食吃不完怎麼辦”的問題。也正因毛澤東受到了媒體誤導，遂於1958年8月的《北戴河會議中》決定大辦人民公社，再次掀起「大躍進」的高潮（董芹，2007）。當時整個國家都在努力實現「大躍進」的高產目標，不論是地方官員或者是媒體，沒有人願意向中央領導承認他們“無法”達成，這也導致各地浮誇風盛行，生產和收成的報告被嚴重誇大（Frank, 2010），這些收成報告經過媒體渲染後傳播給民眾，讓全國人民誤以為當前的國家正在迎來巨大的成功，從而讓「大躍進」的錯誤不斷加劇。

媒體編造大量美麗的神話，使得全國人民充滿幻想，成功的營造出虛假的輿論氛圍，決策者們從報導上看到的也是全國發展一片大好的形勢，因此誤以為制訂的策略是正確的。然而一個地區的上報的產量越誇張，衛星放得越多，該地區的幹部就必須要在中央徵購糧食時承擔相應的代價（董傳嶺，2011；楊繼繩，2009）。最終幹部為了完成徵糧任務，只能不擇手段的從農民身上盡可能地掠奪所有糧食。Xu 和 Tian (2020)認為幹部浮報產量是一種政治忠誠度的展現，且這種展現容易影響鄰近地區的幹部進行相同的行為，若是一地區的浮報行為越氾濫則代表該地區有較強烈的政治狂熱。另外，Xu 等人（2016）曾為了將「各地區高產新聞發佈數量」作為工具變數（Instrumental Variable）取代具有內生性（Endogeneity）的解釋變數：「各地區饑荒嚴重程度」，故隨機抽取《人民日報》中 558 則糧食高產新聞並以省為單位編碼，再對這些樣本進行二階最小平方法迴歸（Two-stage least squares estimator）。結果發現各地區的糧食高產新聞數量對饑荒嚴重程度有顯著的正向影響。Meng 等人（2015）認為「饑荒」之所以持續這麼



長時間，是因為當時媒體報喜不報憂，因此饑荒的事實沒有被宣傳，所以民眾普遍不相信饑荒存在。且當時全國上下壟罩著一股浮誇風，民眾更不可能認為會有饑荒發生，且饑荒的嚴重程度有很大的城鄉差距，居住在城市裡的人並沒有像偏鄉地區的農民一樣面臨如此嚴重的食物短缺。

由上述文獻回顧可以得知，浮誇風的氾濫與大饑荒息息相關，且地區與地區之間的情況亦有所不同。Xu 等人 (2016) 所做之研究與本研究相似性較高，欲驗證之問題也十分雷同，且得出的結果基本上能夠驗證本研究之假設為真。但《人民日報》在「大躍進」進期間所發佈之新聞高達十萬則，其中浮誇風的糧食高產新聞經過本研究初步計算後得知約為 8,000 則以上，而 Xu 等人僅抽樣 558 則做為樣本進行迴歸分析，其結果必然會有較大的抽樣誤差。而本研究利用大數據分析之技術進行問題驗證，能夠有效的減少誤差，得到更精確之結果。除了能夠得知浮誇風報導與饑荒嚴重程度之相關性，且使用糧食徵購量作為中介變數進行檢定，更可以得出全面完整的結果，補足過往研究之不足。

第四節 文字探勘之技術與應用

過往的歷史政策研究通常會使用問卷調查、訪談或焦點小組討論來獲取反饋資訊。然而，這種方法具有針對性，只對特定任務或議題有用，同時也存在樣本收集不全、問卷設計有主觀偏誤等問題 (Ning et al., 2012)。對於國家政策這種範圍較大的議題，小規模調查則缺乏代表性。若要以問卷調查或訪談法進行研究，則需耗費相對大量的人力物力且效率低。因此，當研究對象與問題無法使用小樣本分析結果解釋時，則需要使用巨量資料進行分析，以免樣本缺乏代表性。近年來，學術界投入了大量資源於大數據資料分析，透過使用資料探勘 (Data Mining) 的技術能夠從大量資料中取得傳統方式無法取得的結果。其中，文字探勘 (Text Mining) 被視為是資料探勘的一環，是指利用計算機技術對大量非結構化數據進行分析並得到有價值的信息 (Kang et al., 2021)。這種技術融合了機器學



習 (Machine Learning)、語言學 (Linguistics)、統計學 (Statistics)、線性代數 (Linear Algebra)、機率論 (Probability Theory) 等學門。傳統統計所處理的資料，基本上屬於結構性 (Structural) 的資料，亦即資料與「量」相關，並以表單格式呈現，具明確的結構。文字探勘不同於一般資料探勘的地方在於其內容都是沒有特定規則的純文字，這些文本的內容是由人類的語言所構建的。因此傳統的技術無法對這些「非結構性」的資料進行解析。在文字探勘的領域中，研究者通常會使用電腦輔助的方式對巨量文本資料進行前處理，節省人工逐一閱讀的時間。過去以文字探勘技術進行研究的學者們常以新聞、網路論壇、權威人士的發言、公聽會內容以及學術論文等語料作為研究標的 (邵軒磊，曾元顯，2018)。文字探勘的研究方法也相當廣泛，包括特徵篩選 (Feature Selection)、情緒分析 (Sentiment Analysis)、詞頻分析 (Word Frequency Analysis)、共現詞分析 (Collocation)、文本分群 (Clustering)、主題模型 (Topic Model) 等 (Nassirtoussi et al., 2014)。

新聞是反映人們日常生活的重要文本載體，具有簡潔性的特點。通過對新聞文本的挖掘，我們可以全面掌握事件的現況、趨勢以及不同利益相關者等各方面的反饋 (Kang et al., 2021)。Guan (2018) 曾經使用文字探勘技術於 2001 年 1 月至 2015 年 12 月，15 年間《人民日報》對中日關係的中文新聞報導進行收集與內容分析，該研究使用關鍵字搜索，篩選出符合“日本”和“中日”等關鍵字條件的文本語料，並限定檢索時間為 2001 年至 2015 年。搜索到符合條件的 7,154 篇新聞文章後，對每一篇關於日本和中日關係的新聞報導使用「衝突」、「合作」和「中立」三種概念進行編碼，最後使用主題分析模型分別對這三種概念的報導進行解析，觀察這些概念在 15 年間的主題變化。Yan (2020) 以 2002 年和 2003 年中國爆發嚴重急性呼吸系統綜合症 (SARS) 期間的報導以及 2019 年和 2020 年爆發的新型冠狀病毒 (COVID-19) 期間的報導作為語料，分析中國媒體報導 COVID-19 和 SARS 疫情的用語，探討外部環境與中國疾病報導中軍事隱喻使用之間的關係。



研究結果表明，當代中國疾病報導常會使用戰爭或軍事相關的詞來描述疫情，如“疫情是人民的敵人”、“我們必須防範疫情的侵略”等。這種結果可能與中國傳統的疾病妖魔化有關。Hsiung 與 Wang (2018)收集了 1958-1962 年刊登在新聞上的「順口溜」進行潛在語意和情感分析，發現「大躍進」當時的民眾為了表達不滿常常創作順口溜來諷刺時事與中央，這項研究同時也透漏了「大躍進」時期民眾對當時現況的不滿。Xu 等人 (2010) 則是把《人民日報》1970 年 1 月 1 日至 2011 年 11 月 30 日有關空氣汙染的每篇文章都進行類別（發佈日期、發佈佈局、報告重點以及信息來源）編碼。透過這些編碼量化 30 年間的空汙新聞，進行橫斷面分析。比對中國歷年的空汙狀況數據後發現，歷年新聞數量與報導重點與當年度的空汙嚴重趨勢有關。Bengtsson (2017)為了探討「大躍進」期間，媒體是否有揭露各地饑荒頻傳的事實，作者建立了與“饑荒”、“飢餓”概念相近的相關詞彙作為詞典，並利用該詞典搜索《人民日報》於「大躍進」期間發佈的新聞，結果發現《人民日報》確實發佈過與各地饑荒有關的報導，但相對稀少，因此得出當時《人民日報》違反客觀的報導態度。

綜合以上案例我們可以發現，使用文字探勘技術於媒體資料，確實可以得到許多傳統質化研究無法取得的成果，過往許多研究都使用《人民日報》作為語料進行研究，說明了該數據具有一定的研究價值，且這些學者們使用的技術與研究成果也有助於本研究之推論與參考。然而使用新聞數據集來進行文字探勘分析「大躍進」議題，相較一般質化研究仍佔少數，故可以凸顯本研究之重要性。本研究將依歷史時間順序對文本進行分析，除了以媒體的角度觀察中央與人民如何受其影響之外，亦會透過回顧過去相關之文獻與參考該時期發生之重大事件與政策以站在人民與政府的角度做出全面性的分析。

第三章 研究設計

第一節 研究方法與架構



本研究將以當時中共中央最大的傳播媒體《人民日報》作為切入點，針對1958年1月至1962年12月之新聞報導內容進行文字探勘，同時結合過去文獻理論依據建立迴歸模型。另外，由於浮誇新聞種類繁多，在「大躍進」時期的浮誇風不僅包含糧產，還包含著煉鋼、煉鐵以及科技和技術等各種面向，以邏輯上來說，糧產的浮誇新聞與饑荒嚴重程度的關聯性應該較大，因此本研究定義之浮誇新聞只限與糧產相關，探討之議題主要為浮誇新聞數量與非正常死亡人數趨勢比較，整體研究流程與架構如圖8所示。

壹、資料前處理

本研究以量化方式來分析巨量文本，故得先將大量的非結構性（Unstructured）資料先轉為結構性（Structured）資料，方能進行後續之統計分析以及視覺化呈現。為了要將文本資料結構化，必須先進行資料前處理。首先把每篇文章進行「斷詞」，亦即把文章以「詞」為單位進行分割，以利後續文本分析之演算法計算詞頻。斷詞套件選擇方面，由於研究樣本為中文文本，因此本研究將分別嘗試由大陸開發專門處理簡體中文斷詞的Jieba套件和台灣中研院為處理中文斷詞所開發之CKIP斷詞系統，比較兩者之斷詞效果優劣。此外，為了使斷詞結果更精準，亦會輔以自定義詞典協助斷詞，該詞典之建置除了參考網路上現有的開源簡體中文用語詞典外，也將透過觀察文本內容以自身領域知識新增字詞，最後逐一檢視斷詞錯誤並予以修正。為了方便後續分析結果之判讀，在斷詞完成後，會將無意義之字詞或連接詞刪除，同時刪除標點符號以及特殊符號。然而，許多無意義字詞仍無法透過上述方式排除，導致文本仍然會存在許多雜訊。為了達到去蕪存菁之效果，本研究將分別嘗試使用TF-IDF和TextRank進行文章的關

關鍵字提取。這些關鍵字有助於後續本研究繪製文字雲圖，與詞頻計算，方便後續本研究利用關鍵字分析方法來了解浮誇新聞的組成架構。



貳、糧產浮誇新聞的篩選

為了探討「浮誇風的糧食高產新聞與饑荒嚴重程度之關聯」，本研究需在約十一萬筆新聞中最大化的找出浮誇新聞文本，為此需要利用程式來輔助文本篩選。當前常用的篩選方法主要有「關鍵詞搜索法」以及「文本相似性計算」，然而經過本研究的觀察，浮誇新聞的主題包含了各種面向，例如：工業、糧產、科研技術、人民公社化以及中央領導的歌功頌德等，除此之外，形容浮誇的語句更是千變萬化。基於上述理由，本研究評估難以建立一組能夠正確篩選出所有糧食浮誇新聞的關鍵詞字典。然而「文本相似性計算」可以解決建立大量關鍵詞字典的問題，其原理主要是利用傳統的統計詞頻方法、維度縮減方法或機器學習演算法來建立模型，並將文章字詞透過模型轉為向量後計算各篇文章之間的相似性 (Cosine Similarity)，藉此來衡量文本之間的關係。本研究礙於時間與人力之限制，經評估後決定採用「文本相似性計算」方法來篩選浮誇新聞。然而在計算相似性之前需要先將每篇新聞轉成向量才能進行後續之數學計算。目前學術界主要使用的文本轉向量方法主要採用機器學習演算法所建立之模型來進行，其中常用的演算法又分別以 Le 與 Mikolov 於 2014 年提出的 Doc2vec 演算法以及 Vaswani 等人於 2017 年所提出的 Transformer 模型最多人使用。

Doc2vec 是 Le 與 Mikolov 基於 Word2vec 模型所衍生出的新模型，是以類神經網路之無監督學習演算法訓練而成，且兩者原理非常相似。其中 Word2vec 是將「字詞」轉換成「詞向量」(Word embeddings)。有了「詞向量」便可以針對文本進行後續各種數學計算與分析 (Mikolov et al., 2013)。Word2vec 所使用的演算法主要有兩種，第一種為 Continuous Bag Of Words，此演算法以文本上下文的字詞來當作模型的輸入訓練，最後輸出模型預測的詞向量。而第二種演算法 Skip-

Gram 則與第一種演算法的訓練方式相反，是以一個字詞來預測該字詞的上下文（詳見，圖 2）。

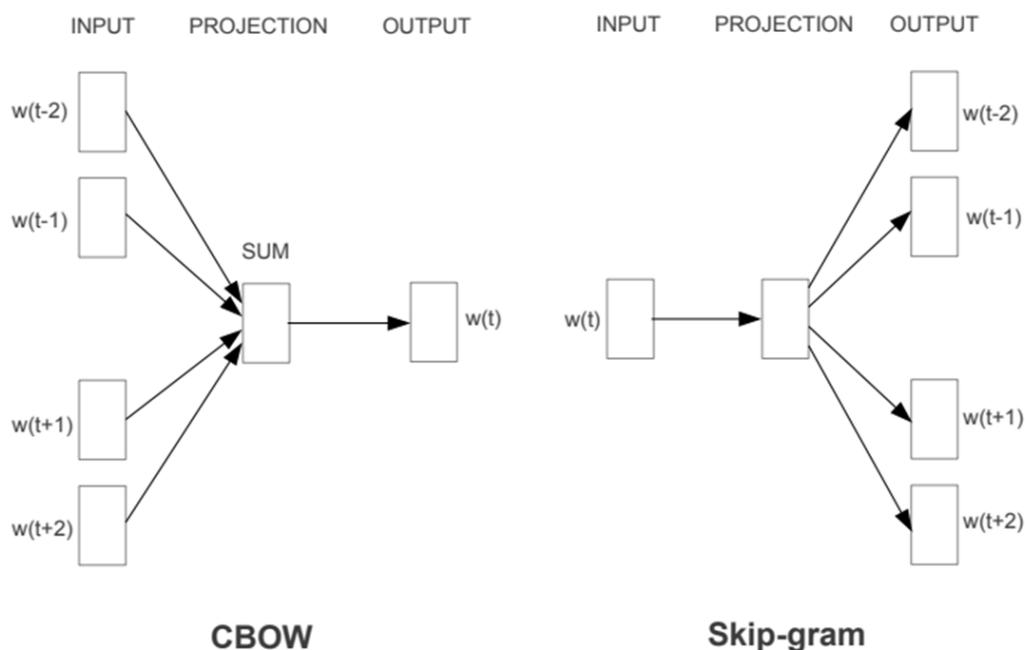


圖 2 CBOW 與 Skip-gram 架構

資料來源：引用 Mikolov et al., 2013 之研究結果

然而 Word2vec 的強項在於轉換字詞的向量，倘若要把一篇文章轉為一個向量則須把文章中所有的字詞轉為「詞向量」後，再將其加總做平均處理，最後以詞向量的平均值代表該篇文章的「文章向量」，這種方法沒有考慮到字詞的先後順序以及前後文之關聯，從而導致無法有效地用向量來表達文章。為了解決這個問題，Word2vec 的作者 Mikolov 等人開發了 Doc2vec 幫助解決模型難以理解字詞上下文的問題，Doc2vec 是一種以 PV-DM (Distributed Memory Model of paragraph vectors) 演算法訓練的深度學習模型，能夠根據輸入詞集計算文章內字詞間的距離，此方法能讓模型考慮到字詞間的前後關係，同時理解文章段落的上下文（詳見，圖 3），能夠更正確的將文本向量化 (Le & Mikolov, 2014)。

Average / Concatenate

Word Embedding

Input

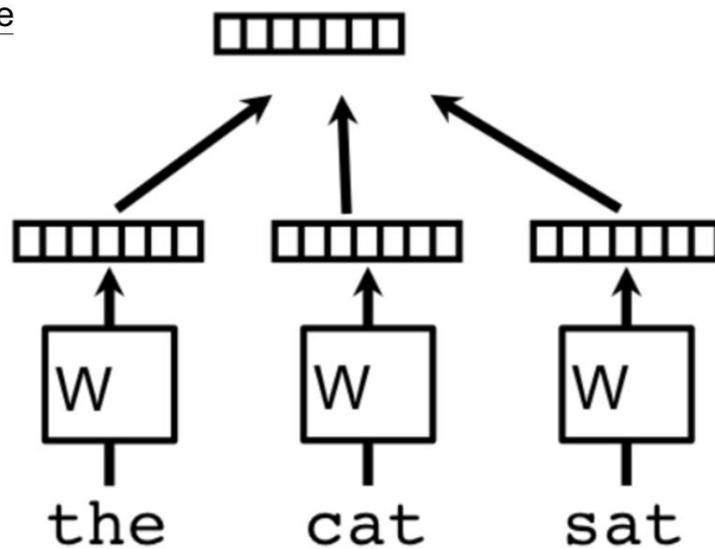


圖 3 PV-DM 架構

資料來源：引用 Mikolov et al., 2014 之研究結果

Transformer 模型為 Google 團隊於 2017 年所提出 (Vaswani et al., 2017)，主要目的是解決 Sequence-to-sequence Models 在訓練時輸入資料經過 Encoder 至 Decoder 輸出的過程中資訊丟失的問題 (詳見，圖 4)。該文提出了新的演算法架構“Multi-Head Attention”讓每一筆序列資料不論距離多遠都能考慮到其前後關係 (詳見，圖 5)。這種模型非常適合處理語言、文本、時間序列等具有前後關係的資料，且在自然語言處理 (NLP) 應用為最大宗，例如機器翻譯、語音辨識、問答系統、字詞向量化等。而利用 Transformer 模型架構所訓練出來的預訓練模型 (Pre-trained Model)，例如：GPT2 與 BERT 除了能夠準確地將字詞轉換為向量外，更能夠完整的表示字詞的上下文語意。

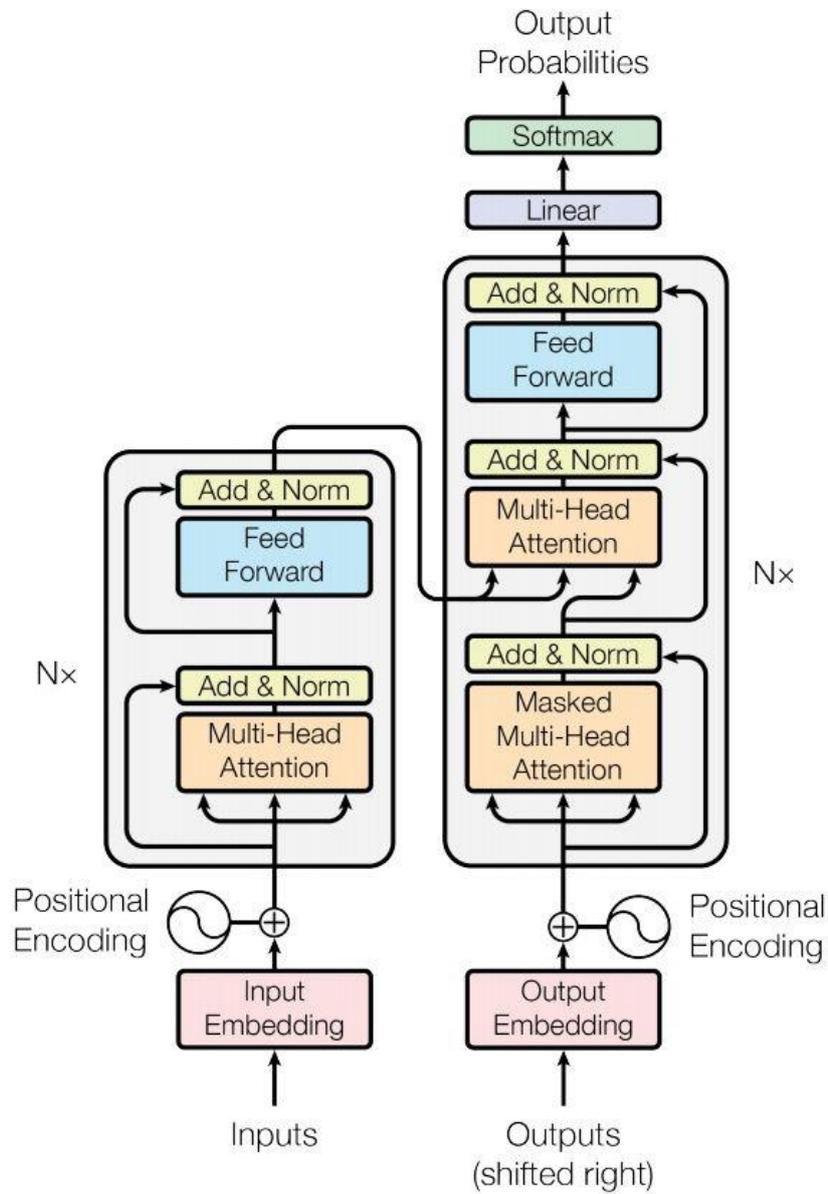


圖 4 Transformer 架構

資料來源：引用 Vaswani et al., 2017 之研究結果

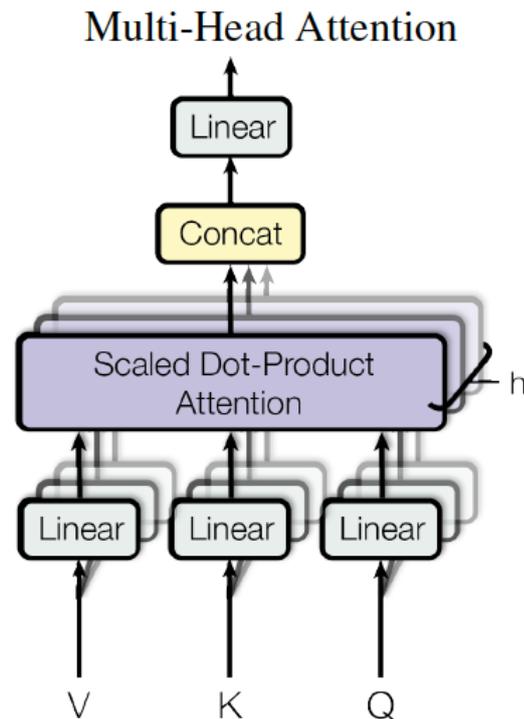


圖 5 Multi-Head Attention 架構

資料來源：引用 Vaswani et al., 2017 之研究結果

本研究嘗試使用 Transformer 的預訓練模型以及 Doc2vec 模型將文本做向量化處理，從文本清理，模型訓練，文本向量化，相似性計算，最後篩選出浮誇文本，流程如圖 6 所示。其中 Transformer 架構所訓練的預訓練模型，經過多方嘗試各種應用於簡體中文的開源模型後發現以 sbert-chinese-general-v2 之預訓練模型所篩選出的浮誇文本效果最為準確，此模型是基於 Google 所釋出的開源模型 Bert 加以校調，在百萬級別的語意數據集上進行訓練，具有良好的泛化能力，且適合本研究之任務。

為了進行文本相似性計算，本研究先挑選出 864 筆確定為浮誇新聞的文本作為目標文本，再利用迴圈的方式將 864 篇目標文本依序讀出，每讀出一篇就分別與所有文本計算相似性。參數設定方面，本研究嘗試以 Doc2vec 將文本轉換為 150 維度、200 維度、250 維度的向量，sbert-chinese-general-v2 模型則預設將文本轉換

為 768 維度的向量，並嘗試設定讓兩模型只取出 Cosine Similarity 大於 0.8 以上之文本。經過測試後發現，Doc2vec 使用 250 維度能夠篩選出 7,631 篇，而 sbert-chinese-general-v2 使用模型預設的 768 維度，可以篩選出 8,772 篇，能在不失準確度的情況下最大程度的篩選出所有浮誇新聞。本研究最後在兩個模型所篩選出的浮誇新聞分別隨機抽樣 200 筆。驗證後發現，sbert-chinese-general-v2 模型準確率約為 86%。而使用 Doc2vec 模型則只有準確率 80% 的效果。經過上述的比較，我們可以得知在文本向量化任務方面，Doc2vec 模型準確率略遜於預訓練模型 sbert-chinese-general-v2，而 sbert-chinese-general-v2 的篩選結果顯示 1958 至 1962 年，人民日報發佈的十一萬則新聞中，大約有 8,772 則與糧食相關的浮誇新聞（如表 6）。

表 6 模型參數設定與篩選結果比較

項目	Doc2vec	sbert-chinese-general-v2
Cosine Similarity 標準	0.8	0.8
文本向量維度（預設）	250	768
浮誇文本篩選數	7,631 則	8,772 則
準確率（抽樣 200 筆）	80%	86%

資料來源：本研究自行繪製

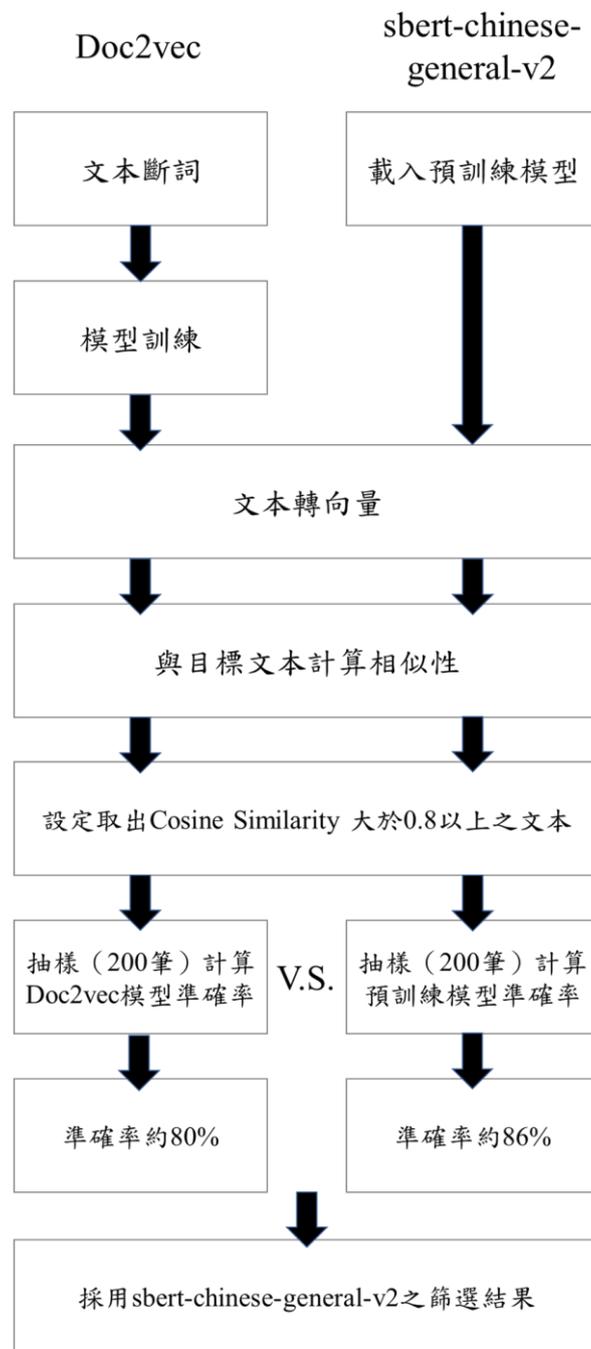


圖 6 文本相似性計算流程

資料來源：本研究自行整理

參、浮誇新聞的關鍵字分析

新聞媒體是如何在當時饑荒的年代塑造出高產榮景實在令人好奇。因此篩選完成後，本研究將以關鍵字分析方法找出浮誇新聞通常都會使用哪些字詞（例

如：亩产万斤、丰产、丰收、史无前例等），透過觀察各個樣本間之語句表達是否存在特定邏輯來了解中共中央官媒《人民日報》是如何形容和創造社會的高產榮景。為了解這些浮誇新聞的組成結構與關注重點，首先須透過特徵篩選把重要的字詞過濾出來後計算詞頻，把出現頻率高且有意義之字詞分別進行累加、排名。最後以高頻重要字詞繪製文字雲圖，判斷浮誇新聞主要的字詞組成與著重的焦點，了解當時的新聞媒體是如何對人民進行思想傳播並成功控制社會輿論。

肆、各省的浮誇新聞數量

本研究後續將以地區（省）為單位分類浮誇新聞，並進行迴歸分析，為此需先將篩選出的浮誇新聞進行地區別的分類，分類作業將使用程式輔助進行。首先透過上一項的文章相似性方法可以大致揭露這五年間與糧食相關的浮誇報導。然而僅使用這種篩選方法無法知道該篇新聞是在報導哪個地區，必須先依照不同「省級行政區」將其分類，因此本項將會在程式的條件判斷式中加入各省之縣、鄉、區等行政區名稱，判斷：「新聞內是否出現省與行政區關鍵字」，利用這種關鍵字搜索來劃分本研究所篩選之 8,772 篇浮誇新聞。舉例來說，若一篇浮誇新聞同時提及了「四川」、「安徽」、「北京」、「上海」，則該篇新聞將會被分別計入這四個省分的新聞發佈篇數，也就是說一篇新聞會因為內容提到之地名多寡而被重複計算至相對應個數的省分，成為以省級行政區為單位的浮誇新聞，並以省為單位呈現加總後的結果（如圖 7）。

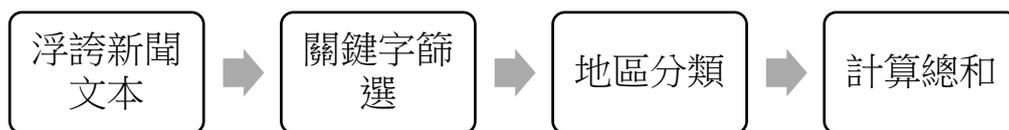


圖 7 地區別浮誇新聞分類流程

資料來源：本研究自行整理



伍、迴歸分析

探討各省的浮誇新聞數量以及非正常死亡人數是為了衡量浮誇風氾濫程度是否會影響饑荒的嚴重程度，同時也是本研究欲探討之最終目的。然而，僅使用各省的浮誇新聞篇數與饑荒非正常死亡人數進行迴歸分析將會忽略不同省份「新聞發佈總數對浮誇新聞數量的影響」，以及「不同地區的人口多寡對非正常死亡人數的影響」。為此本研究將會把「大躍進」五年間《人民日報》發佈的 112,484 篇新聞比照前一項浮誇新聞地區別分類方法，計算出各省份的新聞發佈總數（包含浮誇新聞與所有議題之新聞），最後將各省「浮誇新聞數」除以「總新聞數」再乘以 100% 得出各省的「浮誇新聞發佈百分比」，藉以控制不同地區總新聞數量對浮誇新聞數量的影響。另外，為了控制不同地區的人口數差異對非正常死亡人數的影響，本研究將各省份的「非正常死亡人數」除以「總人口」再乘以 100%，得出各省的「人均非正常死亡百分比」。讓解釋變數與被解釋變數皆以百分比為單位，不僅可以控制其他因素帶來的影響，也能方便後續迴歸分析結果之判讀。除此之外，本研究也將繪出「人均非正常死亡百分比」（縱軸）和「浮誇新聞百分比」（橫軸）等二維度散佈圖並加上迴歸預測線。把圖中的所有樣本點分別標記地區名稱，如此即可辨別出省與省之間浮誇新聞與非正常死亡之差異，顯示出當年國內發生饑荒的 28 個省級行政區於兩個維度之分佈情況。透過樣本點與迴歸線的相對位置，也可以得知哪些省份的死亡人數被高估或低估。除了驗證研究假設是否成立之外，與迴歸線距離相去甚遠的省份我們也將另外個別檢視這種歧異值現象的可能原因為何，並補充說明之。

本研究的饑荒嚴重程度之衡量依據會以楊繼繩（2009）所推估之各省歷年非正常死亡人口作為對照的數據，由於楊繼繩的資料是統整官方資料與參考各學者之計算方式推估而來，儘管相較於其他學者提出的結果更為保守，但其研究抱持著對官方數據的存疑，並利用各地的機密文件來揭開「大饑荒」死亡的真實面

貌，廣受學界大部分學者所推崇，且其資料更完整的提供各省的歷年非正常死亡人口。

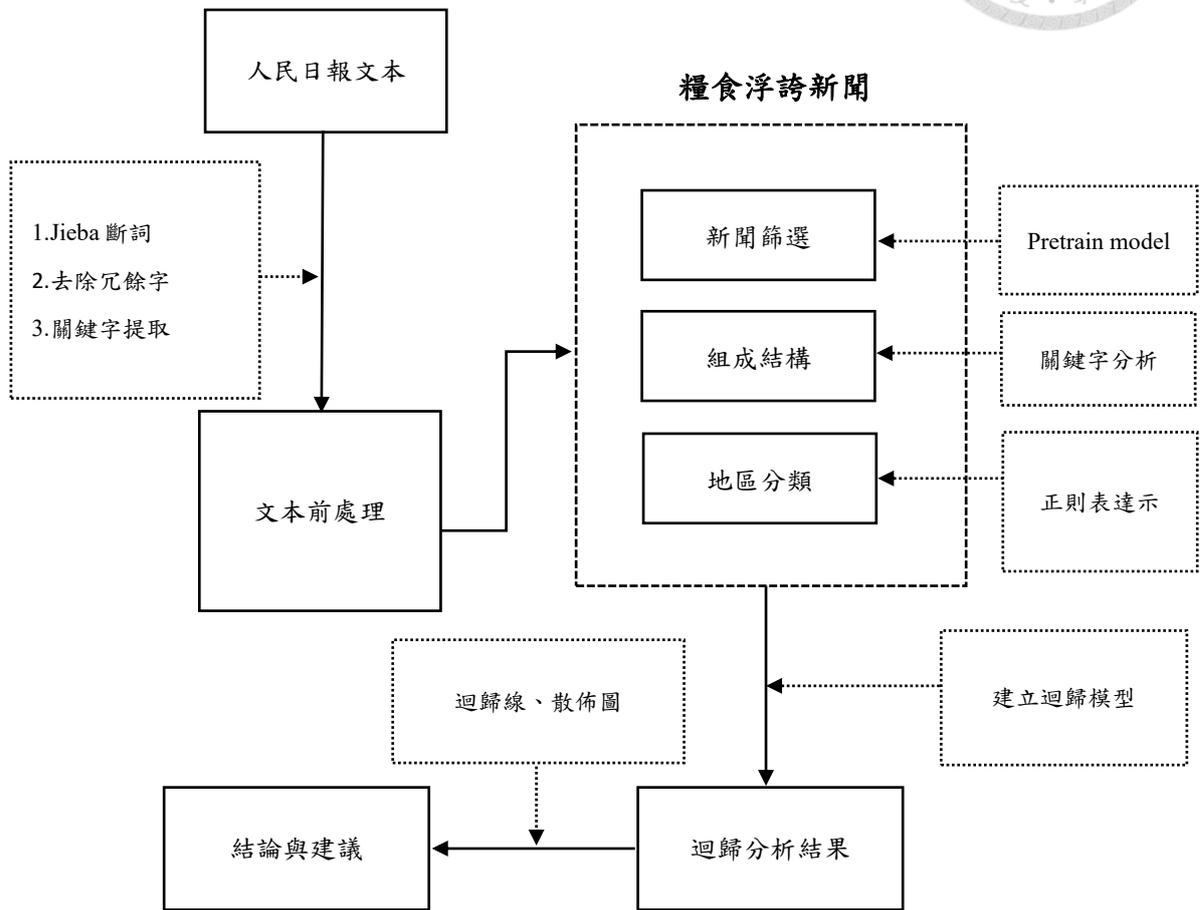


圖 8 研究流程與架構

資料來源：本研究自行繪製

第二節 研究對象與資料來源

壹、各分析方法之研究對象

本研究之分析標的為「大躍進」時期中共中央媒體《人民日報》所發佈之新聞報導。過去研究中共媒體之相關文獻也經常使用《人民日報》的新聞作為研究樣本，其主要原因在於《人民日報》之歷史地位以及對社會之影響力巨大。儘管



「大躍進」時期，發佈新聞的媒體除了《人民日報》外，還有新華社、《紅旗》黨刊雜誌，以及隸屬於中宣部的《光明日報》，但是身為中央黨報的《人民日報》不論是規模、政治背景或是言論的影響力都屬其中最大，因此以《人民日報》的新聞探討當時的媒體之行為具有代表性。研究範圍為設定在 1958 年 1 月至 1962 年 12 月，也就是「大躍進」持續的五年間《人民日報》總共發佈了 112,484 篇新聞報導。

貳、資料來源

本研究之資料來源為「人民日報圖文數據庫」，該資料庫紀錄了《人民日報》1946 年 6 月創刊至今之新聞報導，亦包括圖文資料與網頁版新聞以及海外版新聞。《人民日報》是中國最具影響力和權威性的報紙，該資料庫直接來源於人民日報社，過去 60 年，《人民日報》不僅作為中共中央向國內與世界發佈政策法規、政治動向、經濟資訊的最大管道，同時也是研究現代中國的學者的重要資源。為了獲取「大躍進」期間之新聞報導，本研究透過台灣大學之 VPN 登入資料庫中，利用台灣大學之身份獲得存取資料之權限。

第三節 分析工具

Python 程式語言在進行大數據分析與文字探勘方面十分便利，網路上亦有許多開源的程式碼與套件可供 Python 使用者免費使用。在資料清理方面將使用 Jieba 斷詞套件進行斷詞；浮誇新聞篩選則會分別使用 Gensim 套件中的 Doc2vec 模型以及 Sentence-transformers 套件中的開源簡體中文預訓練模型嘗試將文本進行向量化處理；浮誇新聞關鍵字分析則會選用文本特徵篩選的 TF-IDF 與 TextRank 模型將關鍵字取出，再以視覺化（文字雲圖）呈現結果；浮誇新聞的地區分類將使用 Python 中內建的 re 模組提供的正規表示法（Regular Expression）針對每篇文本進行地區關鍵字的掃描，以省份為單位累積浮誇新聞。經過層層篩選與計算後，加

總「大躍進」五年期間各個省份所累計的浮誇新聞總數。迴歸分析會使用 python 開源的 Stats models 套件來執行。最後視覺化的迴歸線散佈圖以及五年間的全國非正常死亡、浮誇新聞、糧食徵購趨勢比對則會以 Matplotlib 套件呈現。



第四節 論文結構與章節安排說明

本論文之章節安排是以問題發想至結論建議之過程順序而設計，一共分為五章。第一章「緒論」，共分為兩節，其中第一節「研究緣起」主要說明本研究探討議題之核心、發想以及價值，透過「研究緣起」，讀者可以初步了解本研究欲探討問題之重點與意涵。第二節「問題意識與主要研究問題」是聚焦於研究議題之介紹，包括該議題之時空背景與後續影響之分析，讓讀者更深入了解議題與本研究該如何著手進行研究。具體定義本研究欲探討之問題面向，界定探討範圍以及該從何種角度切入探討，最後列出要回答的問題有哪些。

第二章「相關理論回顧與文獻探討」，共分為四節，每節包含兩項：第一節為「大躍進與大饑荒相關研究」，探討之項目為過去學者如何看待「大躍進的發起與影響」以及「三年大饑荒的成因與死亡人數」。第二節「中國大陸媒體之研究」主要以「黨與媒體的關係」和「大躍進時期媒體之傳播途徑」這兩項目之角度來回顧過去學者對中國大陸媒體扮演之角色與行為以及媒體發出的訊息從何種途徑進行傳播進而影響群眾。第三節為「浮誇風與饑荒的關聯」，為本研究最重要之文獻回顧，透過檢視過往與本研究目的雷同的相關研究，確認研究假設之合理性與正確性。第四節為「文字探勘之技術與應用」，分別回顧文字探勘之技術原理與相關應用以及參考過去使用文字探勘進行與本研究議題類似之方法與技術。

第三章「研究設計」闡述本論文之研究流程、分析方法、樣本來源與整體研究架構。分別為第一節「研究方法與架構」：說明本研究之分析方法、整體設計、衡量項目切入面向以及執行流程；第二節「研究對象與資料來源」詳述各分

析方法之研究對象，說明研究樣本的描述性統計以及資料來源、獲取方式；第三節「分析工具」說明本研究之分析以及資料處理分別使用了哪些工具，以及選用之理由。第四節「論文結構與章節安排說明」目的為統整論文結構和章節安排。

第四章「研究結果」為論文之正文章節。編排將依據研究目的中的問題並分為三節，分別為「浮誇新聞的篩選與分類結果」、「浮誇新聞的關鍵字分析」以及「迴歸分析」，詳述分析過程與結果。除此之外，引用第二章「相關理論回顧與文獻探討」之理論依據與史料進行解釋。

第五章「結論」為論文各章研究發現之彙整以及後續研究建議。主要分為兩節：第一節「主要研究發現」為第四章結論之彙整，整理分析的結果，回應問題意識，並對各省浮誇風現象和饑荒嚴重程度之衡量結果做最後總結。第二節「後續研究建議」則是根據第一節之結論梳理出本研究有哪些不足之處，建議後續進行相關研究的學者可以往哪個方向接續研究，同時提出有哪些議題與本研究有關且值得探究，供未來研究者參考。

第四章 研究結果



第一節 浮誇新聞的篩選與分類結果

壹、資料前處理

本研究之文本資料來自《人民日報》官方資料庫，1958年至1962年發佈之新聞報導一共有112,484篇，經過初步的觀察後發現，原始文本中仍存在許多亂碼、空白欄位、純圖片資料等雜訊。因此在資料清理方面，本研究將使用Python中的條件判斷式把有問題的文本予以刪除，以確保後續分析之準確性。經過一系列的資料前處理後，能用於分析的乾淨樣本約為十一萬篇。在自然語言處理(NLP)的領域中，為了讓程式的演算法能夠有效辨識文本，通常會先對文本進行「斷詞」處理，以方便後續更進一步分析。所謂「斷詞」是利用模型將一個句子或一篇文章所分割成數個以空白格開的「有意義詞彙」。因此我們可以利用「斷詞」的結果來計算「有意義詞彙」出現的頻率、次數，從而將文字進行做量化處理。斷詞模型的準確率與研究後續分析的正確性息息相關。其中，中文語句的斷詞相較於英文困難許多，以英文語句的結構來說，由於字詞之間是以空格連結，只要以空格為分割依據即可準確區分出正確字詞。因此，中文與英文斷詞的最大差別在於中文有歧異問題。舉例來說，利用斷詞模型將「中國國民黨」進行斷詞，其結果可能為「中國國民黨」、「中國/國民黨」、「中國國民/黨」，可以明顯看出，一個句子可能有數種斷詞結果。這時可能有幾種辦法解決，例如：使用預設詞庫（使用斷詞套件提供的預設詞庫增加斷詞準確性）、自定義字典（利用人工自定義的字典直接限制模型針對特定字詞的斷詞結果）、斷詞模型選擇（選擇專門處理與目標文本相似性質的模型，如Jieba套件擅長處理簡體中文文本，CKIP斷詞系統則擅長處理繁體中文文本）。

由於《人民日報》屬於簡體中文文本，而經本研究分別嘗試Jieba與CKIP兩

套件之斷詞效果後，最後評估斷詞的正確性、完整性因素而選擇以 Jieba 套件所提供之斷詞模型得出的結果作為本研究後續使用之文本。同時也使用了開源的自定義字典來增加斷詞的準確性，最後將 1958 至 1962 年《人民日報》所發佈之十一萬篇新聞分別斷為以空白格連接字詞的資料集。

貳、浮誇新聞之表述方法

經本研究觀察，浮誇新聞大致可分為兩類，第一類為簡要形容全國性的高產盛況，主要目的在於建立「大躍進」政策順利進行的假象以及展現對中央領導的忠誠，這類新聞的範例如表 7 所示。

表 7 第一類浮誇新聞簡要範例

No.	新聞範例
1	<u>大搞群众运动冬季生产上马扬鞭秋粮超夏粮已成定局</u> ，超包产运动全面发展人民公社社员，意气风发，为争取明年农业生产更大跃进。
2	<u>力争多种</u> ，是争取今年农业丰收的重要环节。今年粮食所以取得较好的收成，一方面是由于党的各项政策贯彻的好，干部作风转变的好，调动了广大社员的生产积极性
3	人民公社运动的胜利，加速了我国农村经济的发展，特别是原来经济发展较慢的生产队，发展速度更是惊人。在 1958 年和 1959 年两年间， <u>每年又都比上一年增长了 10%左右</u> ， <u>这样的速度在我国农村经济发展史上是空前的</u> ，世界任何资本主义国家的农村经济发展史上也不曾有过。
4	“试验田”，这个大跃进中诞生的马克思主义的领导方法，现在， <u>这个领导方法又在争取生产大幅度增长中发挥威力了</u> 。这个群众路线的领导方法，在 1958、1959 年的大跃进中已经放出了闪亮的光彩。

5	两年来的大跃进，是毛泽东思想的胜利，是党的总路线的胜利，也是大搞群众运动的胜利群众路线， <u>农业战线，苦战一冬，开荒三百八十三万亩，等于解放十年来开荒总和的二点七倍，等于 1959 年全省播种面积的一半以上；兴修水利可扩大灌溉面积二百八十万亩，等于十年来扩大灌溉面积的总和。</u>
6	在完成新民主主义和社会主义的革命之后，立即展开规模宏大的社会主义各方面的建设。近年来各省市、自治区对于各种农作物的地方品种，因此，在最近粮食征购工作中， <u>很多干部和群众又讨论起要不要向国家交售一些余粮的问题。</u>
7	夏季就已经完成了全年粮食征购任务，秋季又卖余粮支援城市，支援灾区，支援国家建设。 <u>今年夏季收成较好，一次就完成了全年的粮食交售任务，秋季又获得了较好收成，队队实产超包产。</u>

資料來源：本研究自行整理

第二類浮誇新聞則佔多數，其特徵為清楚描述「某地區」的高產盛況、生產成果數據，或以編造故事的方式來讚頌當地的表現，和第一類的差別主要在於，除了建立社會的高產榮景外，同時也反映了各地區的幹部受制於當時中央下達的產量目標指示所帶來的壓力（如表 8 所示）。另外，第二類新聞除了佔浮誇新聞的多數外也是本研究欲分析之重點，因為其記錄了浮報產量與假新聞發生的地區，因此能反映出各地的政治狂熱程度。從表 8 可以看出，《人民日報》在針對浮誇新聞的描寫，通常會一次報導多個地區的高產盛況，也就是說一篇新聞可能會被記入數個不同地區的浮報次數。

表 8 第二類浮誇新聞簡要範例

No.	新聞範例
1	<u>河南省已积各种肥料四千多亿斤，达原计划的 70%，超过 1956 年同期积肥数量四倍以上。甘肃省积肥已达到需肥量的 60%，不仅数量多，而且质量好。著名的高产粮区广东的潮安、澄海、揭阳、普宁、潮阳等五个千斤县，正以“千船万载”的气魄开展冬季积肥运动。</u>

2	<p>责辽宁生猪家禽数量逐月上升河北深县两个生产队大力帮助社员养猪新华社沈阳30日电辽宁省猪、鸡数量逐月上升。据不完全统计，到4月末止，全省生猪头数已达三百五十万头，比去年年底增加了11%，鸡已发展到一千万只，比去年增加了43%；各地还出现了一些养猪、养禽的先进市、县、公社和生产队。开原县现在饲养的生猪头数，已比去年年底增加了61%，抚顺市去年年底只有家禽四万只，4月份已发展到四十多万只。</p>
3	<p>新华社杭州30日电浙江省广大社员努力多种和种好油菜。到10月20日，全省已播种育秧十一万多亩，主要产区嘉兴、宁波、温州等专区已基本种完，许多油菜秧田一片嫩绿，生长茁壮。浙江是我国油菜主要产区之一。今年各地根据国家、集体、社员三方面的需要，因地制宜地扩大油菜种植面积。主要产区平湖、海宁、嘉兴、慈溪等十多个县的人民公社，安排落实种植油菜的面积普遍比去年增加。</p>
4	<p>隆冬天气，西沟大队一片热气腾腾社员竞比高：看来年生产准备谁最好李顺达申纪兰带头劳动，社员积极送粪修农具垒梯田垫河滩据新华社太原30日电一个“看谁为1962年生产准备得最好”的冬季生产竞赛，正在山西平顺县金星人民公社西沟生产大队轰轰烈烈地展开。西沟生产大队是全国著名农业劳动模范李顺达和申纪兰领导的。</p>
5	<p>四川开展油菜播种育苗工作四川许多地区开展油菜播种育苗工作。目前，达县、温江、南充、万县等六个专区已播种油菜十五万多亩。在播种育苗工作中，各地一般都建立了责任制度，保证了进度和质量。</p>
6	<p>坚持“三包一奖”合同省庄大队夏季预分奖励政策全部兑现超产队争取秋季更多超产，赔产队决心迎头赶上本报南昌30日电江西新建县望城岗公社省庄大队在夏收预分中，切实依靠群众，全面贯彻奖励政策，进一步调动了社员的生产积极性。省庄大队今年早稻获得了较好收成。</p>
7	<p>抓季节产酥油青海牧区各公社正抓紧产奶旺季，生产酥油、曲拉（酥油的副产品）。果洛藏族自治州六个县已产酥油六十七万多斤。修造秋收农具福建闽东地区手工业部门及早动手修制秋收农具，支援农业。</p>

資料來源：本研究自行整理



參、新聞地區別分類方法之定義與結果

本研究使用 sbert-chinese-general-v2 預訓練模型，透過文本相似性計算，在 1958 至 1962 年人民日報發佈的十一萬則新聞中篩選出了 8,772 則與糧食相關的浮誇新聞。為了分類這些浮誇新聞的地區來源，本研究建立一組以「區」為單位的地名字典，並依照研究設計章節之分類定義，但凡該篇新聞出現字典內的「地名」，則會將該篇新聞計入該「省」的發佈數。最後統計發現「大躍進」五年期間各地區的新聞數量總和共有 31,014 則，結果如表 9 所示。

表 9 「大躍進」五年期間各省浮誇新聞數量統計結果

省份	次數	省份	次數	省份	次數	省份	次數
北京	359	山東	1,872	黑龍江	1,548	雲南	871
天津	298	河南	1,732	上海	296	陝西	937
河北	1,579	湖北	1,199	江蘇	1,288	甘肅	1,963
山西	905	廣東	1,433	浙江	744	青海	876
內蒙	392	廣西	1,923	安徽	1,005	新疆	219
遼寧	1,286	四川	2,389	福建	1,069	湖南	1,736
吉林	996	貴州	704	江西	1,288	寧夏	309

資料來源：本研究自行整理

本研究也使用 1958 至 1962 年各省浮誇新聞發佈數量繪出歷年之時間變化趨勢。可以從圖 9 可以明顯看出，以全國來說浮誇新聞、非正常死亡，在各年的變化趨勢基本一致（從 1958 年開始上漲，至 1960 年達最高峰後開始下跌），此種漲跌趨勢的結果與「大躍進」政策的執行進程相符，因此可以初步判定浮誇新聞、非正常死亡數具有相關性。

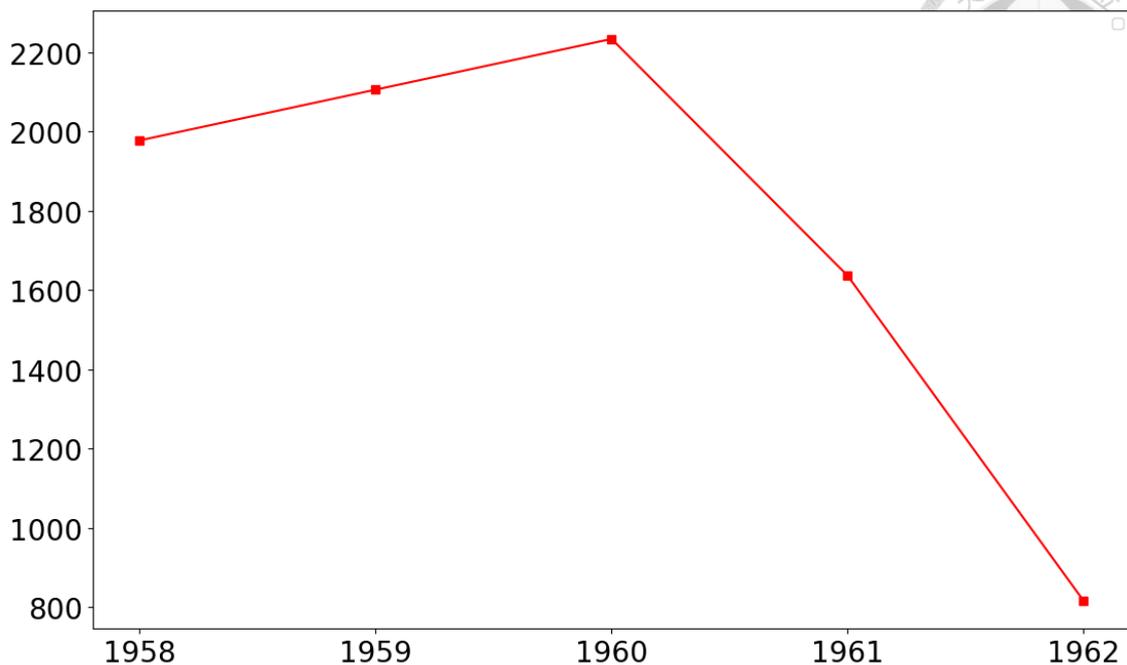


圖 9 歷年浮誇新聞數折線圖

資料來源：本研究自行繪製

第二節 浮誇新聞的關鍵字分析

壹、關鍵字提取

在前一節中，每篇文章已經完成了斷詞的動作，然而就算使用斷詞的自訂義字典或是去除停用詞等方法，整體文本還是存在著相當多的雜訊。因此，想要達到文本去蕪存菁的效果，就必須把整篇文章中最具代表性的數個關鍵字萃取出來。故本研究嘗試使用 TF-IDF (Term Frequency - Inverted Document Frequency) 以及 TextRank 演算法來進行關鍵字提取工作，經過人工主觀判定後發現 TF-IDF 為樣本提取出的關鍵字較為完整且具代表性。因此本研究最終決定以 TF-IDF 作為關鍵字提取的演算法。TF-IDF 主要包含兩個部分：詞頻 (Term Frequency) 與逆向文件頻率 (Inverse Document Frequency)，主要目的是針對文字進行分析，計算字詞的重要性並賦予權重。「詞頻」(TF) 代表某字詞出現在一篇文本的頻率，計算方式為一文內該詞數量除以總字詞數。若該字詞的出現次數越多則 TF 值越高，

代表字詞在這篇文章中的重要性越大。然而有許多字詞的詞頻非常高但卻不具代表性，例如連接詞、語助詞等冗餘字。因此只用詞頻來衡量字詞在文本中的重要程度明顯不夠嚴謹。為此則需要搭配「逆向文件頻率」幫助衡量。所謂逆向文件頻率（IDF）為計算字詞出現在數據集內的幾篇文本中，具體計算方式為總文本數除以出現過的文本數，若該字詞出現在越少文本內則越重要，IDF 值也越高。得知 TF 值與 IDF 值後再將兩者相乘，即為 TF-IDF 值，該值能更準確的衡量一個字詞的重要性與代表性。研究者可以設定要保留多少 TF-IDF 值以上的字詞作為文本的代表字詞，這些字詞同時也代表該文本的特徵，除了可以更準確的進行後續分析，也可以大量減輕程式的運算負擔。

貳、關鍵字頻分析

每篇新聞經過了 TF-IDF 加權後，大部分冗餘字已經被刪除，留下的字詞皆為能夠代表文本的關鍵字組。有了各篇文本的關鍵字，本研究即著手累積各篇文本字詞的出現頻率。由出現頻率排名前二十的字詞（詳見，表 10）可判斷出《人民日報》在撰寫糧食相關的浮誇新聞時，主要著重在「人民公社」、「豐收」、「大躍進」等相關字詞，這些字詞明顯無法看出當時國民正在飽受饑荒之苦。

表 10 關鍵字出現頻率前 20 名（單位：次）

排名	關鍵字	詞頻	排名	關鍵字	詞頻
1	人民公社	2,884	11	发展	1,603
2	生产队	2,461	12	作物	1,599
3	增产	2,290	13	劳力	1,565
4	万亩	2,052	14	面积	1,546
5	组织	2,041	15	平均	1,522
6	专区	1,937	16	保证	1,496
7	粮食	1,689	17	积肥	1,481
8	提高	1,683	18	丰收	1,366

9	产量	1,660	19	每亩	1,315
10	播种	1,622	20	大跃进	1,301

資料來源：本研究自行整理

本研究亦找出了前 35 個浮誇新聞經常用來形容農產大豐收的字詞，如：「農業高潮、規模壯闊、生產競賽、不斷提高、好几倍、大規模、高額、空前、新紀錄、大發展、大規模、史无前例、衛星、榜樣、奇蹟、大量、史无前例、奇蹟、新紀錄、創紀錄、高產、創新高、最高紀錄、扶搖直上、創舉、振翅欲飛、又增加了、全面躍進、躍進浪潮、超產、大增產、田滿、滿栽、滿插、穗大」等等。透過計算這些浮誇字詞分別在 1958、1959、1960、1961、1962 年的《人民日報》報導中出現的次數可以看出，浮誇字詞於《人民日報》的聲量變化是由 1958 年開始逐步攀升，一直到 1960 年達到最高峰並開始急遽驟降（圖 10 所示）。

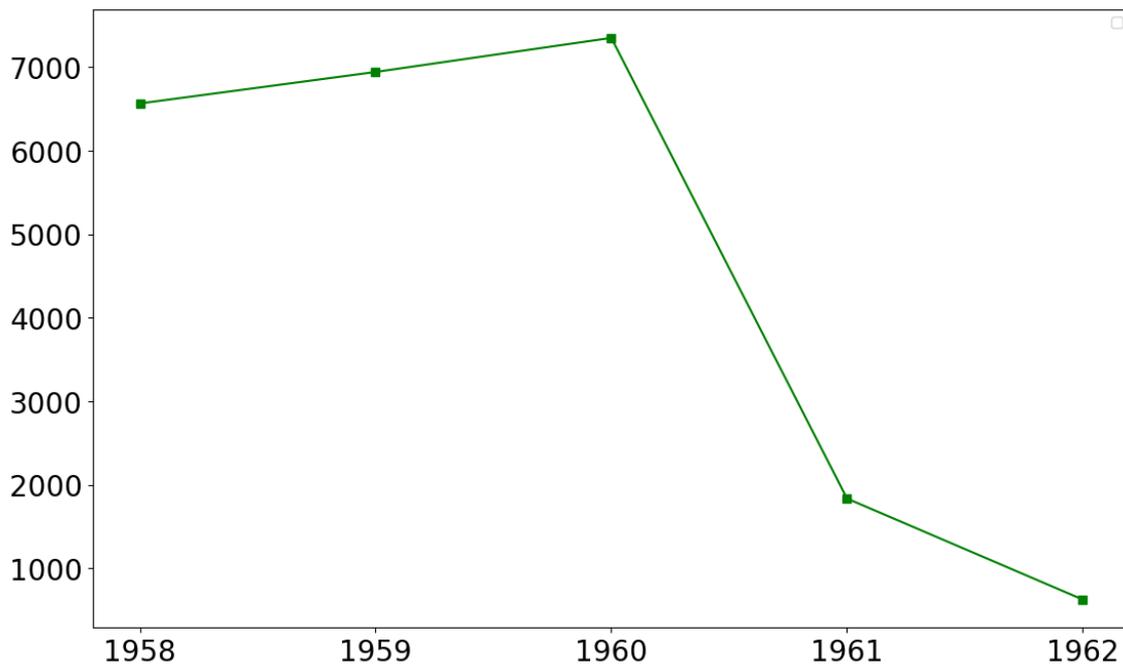


圖 10 關鍵字聲量折線圖

資料來源：本研究自行繪製

除此之外本研究也使用出現頻率排名前 200 的關鍵字來繪製文字雲圖（如圖 11），並依據出現次數的多寡來決定字詞在文字雲圖中出現的大小。可以發現這些高頻字詞包含許多形容各地豐收與糧食充沛的形容詞，例如「豐收」、「高

第三節 迴歸分析



壹、樣本敘述統計

本研究利用程式篩選所得出的各地區浮誇新聞總數為 31,014 則，然而為了控制不同地區總新聞數量對浮誇新聞數量的影響以及不同地區的人口數差異對非正常死亡人數的影響。本研究透過將各省「浮誇新聞數」除以「總新聞數」以及各省「非正常死亡人數」除以「總人口」並分別再乘以 100%，得出「浮誇新聞數百分比」與「人均非正常死亡數百分比」。從樣本的敘述統計（如表 11）可以發現，在控制各省的總新聞發佈數量後，浮誇新聞發佈比率最少的地區為北京 1.96%，最多的地區為湖南的 25.1%，整體樣本的中位數為 12.88%，平均值為 13.09%，標準差為 6.44%。各省非正常死亡人數在控制總人口後可以得知人均死亡比率最少的地區為內蒙 1.96%，最多的地區為四川 2.38%，整體樣本的中位數為 0.27%，平均值為 0.46%，標準差為 0.52%。

表11 樣本敘述統計表

N: 28	最小值	最大值	中位數	平均值	標準差
浮誇新聞數 (%)	1.96	25.1	12.88	13.09	6.44
人均非正常死亡數 (%)	0.02	2.38	0.27	0.46	0.52

資料來源：本研究自行繪製

為完整瞭解研究樣本，本研究將為「浮誇新聞數百分比」、「人均非正常死亡數百分比」等變數繪製盒鬚圖，觀察變數的最大值、最小值、中位數、第一四分位數 (Q1)、第三四分位數 (Q3)，以及四分位距 (Interquartile range, IQR) 之分佈，同時觀察變數中是否存在離群值 (Outlier)。如果樣本距 Q1 (第一四分位數) 或 Q3 (第三四分位數) 過遠 (本研究定義為超過 Q3 加上 1.5 倍的 IQR 或低

於 Q1 減去 1.5 倍的 IQR)，則將樣本視為離群值。圖 12 為「浮誇新聞數百分比」與「人均非正常死亡數百分比」之盒鬚圖，可以看出「浮誇新聞數百分比」並沒有樣本為離群值，但「人均非正常死亡數百分比」則有「四川」為離群值，可以得知四川的非正常死亡比率遠超平均許多。

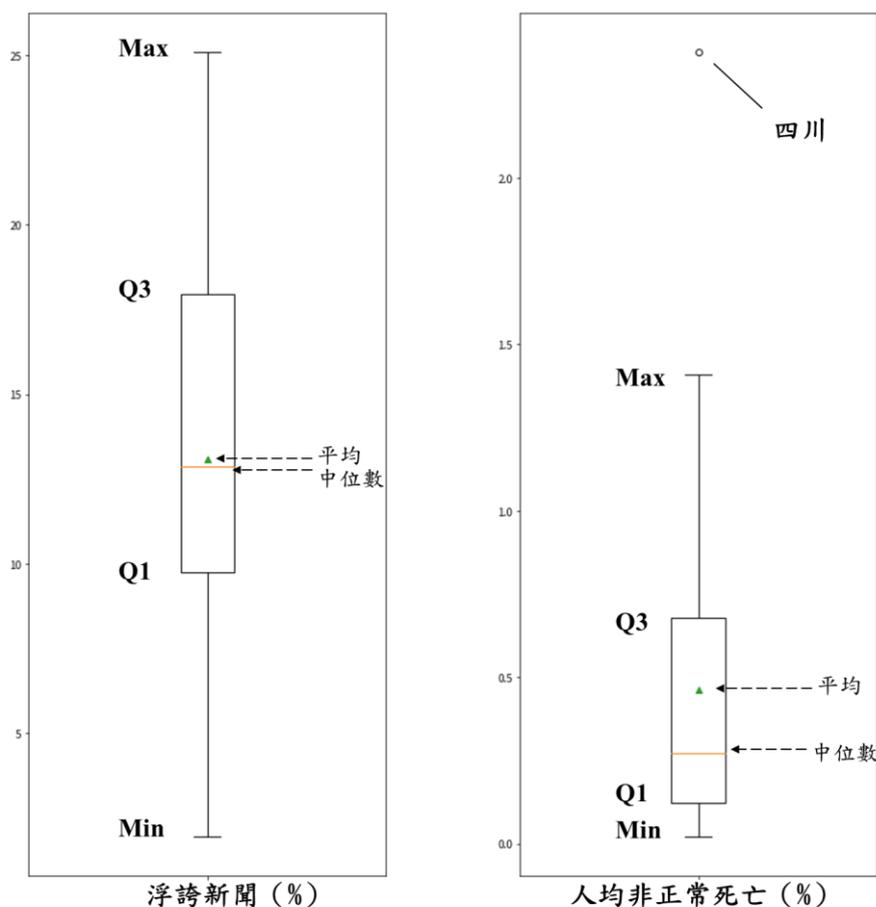


圖 12 「浮誇新聞數百分比」與「人均非正常死亡百分比」盒鬚圖

資料來源：本研究自行繪製

本研究考慮到城鄉地區饑荒程度的差距，因此將樣本區分成「自治區」、「省」、「直轄市」三種類別，計算出這三種類別之「人均非正常死亡數百分比」平均數，以及「浮誇新聞數百分比」之平均數，再分別繪出長條圖。可以明顯發現直轄市的「人均非正常死亡數百分比」相對於省與自治區少（如圖 13），可能原因為直轄市並非「大躍進」時期的農業重鎮，因此糧產壓力相對較少，且當時

的重點農業地區（如四川、安徽）甚至常常將糧食外調至發達的直轄市（北京、天津、上海），這也導致了一些農業省份的饑荒越來越嚴重，而直轄市卻沒有太多非正常死亡發生。而省與自治區的「浮誇新聞數百分比」平均數差異不大（如圖 14），但可以看出兩者與直轄市仍有較大差異，這種現象其實也與該地區是否為農業城市有關，因為本研究定義的浮誇新聞主要是在形容糧食的高產量，因此非農業重鎮的直轄市自然也不會發佈太多浮誇新聞。

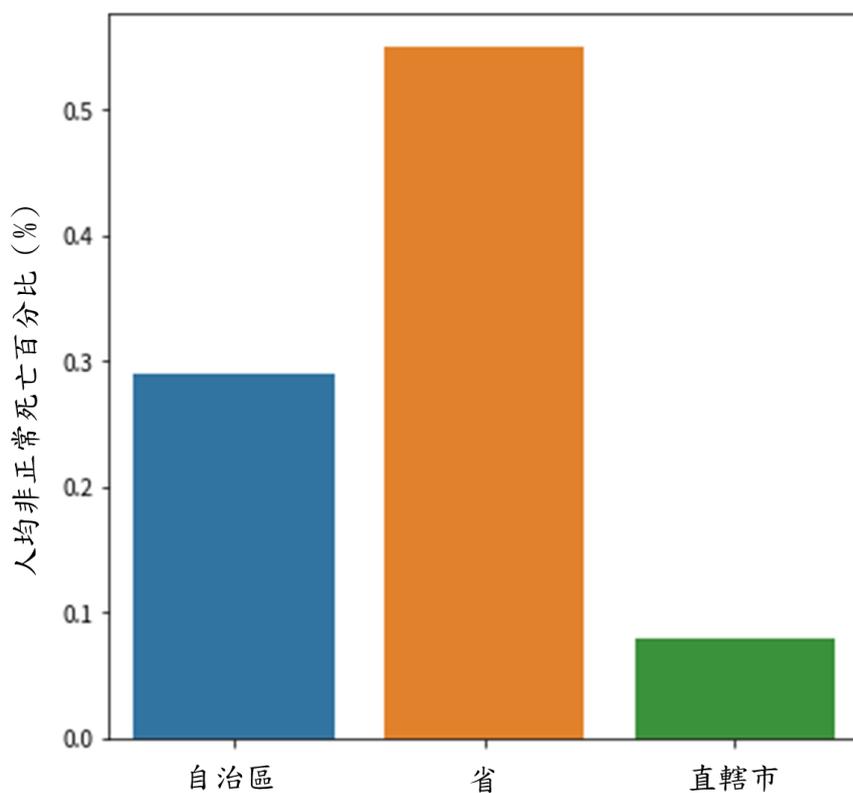


圖 13 人均非正常死亡數百分比長條圖

資料來源：本研究自行繪製

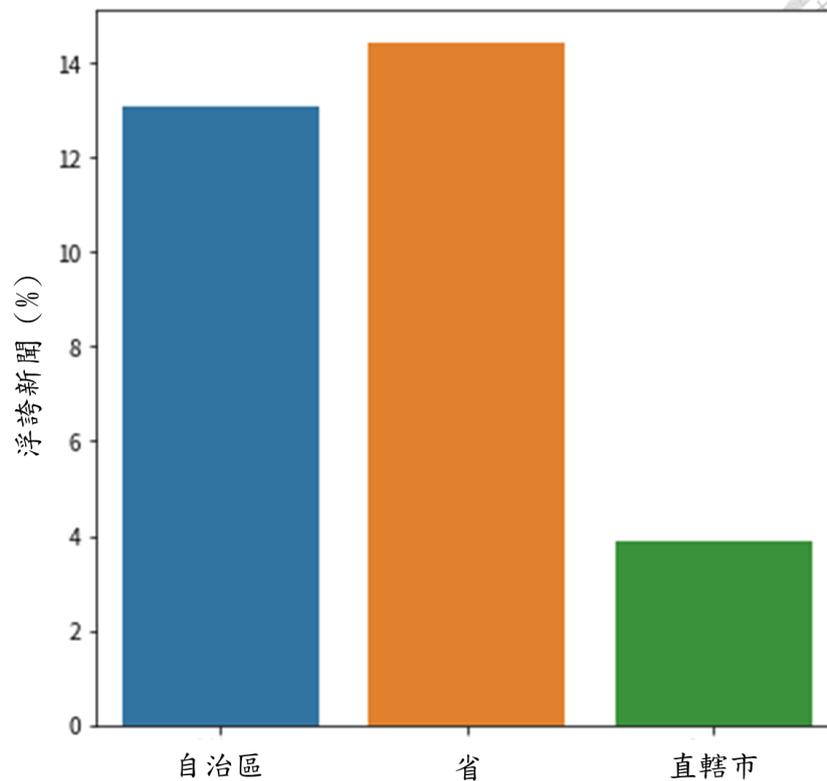


圖 14 浮誇新聞數百分比長條圖

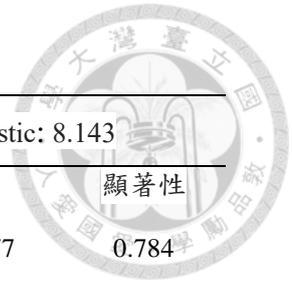
資料來源：本研究自行繪製

貳、分析結果

本研究以「大躍進」期間饑荒發生的 28 個省級行政區作為分析單位來建構迴歸模型，並以「浮誇新聞數百分比」作為解釋變數，「人均非正常死亡數百分比」作為被解釋變數進行迴歸分析。研究結果顯示（如表 12），「浮誇新聞數百分比」的迴歸係數為 0.04，P 值為 0.008，模型的 R^2 為 0.239，此結果代表「浮誇新聞」對「人均非正常死亡」有顯著的正向影響，即研究假設成立。本研究亦繪製出所有地區之樣本點分佈之散佈圖並加上迴歸預測線（如圖 15）。從迴歸線呈現正斜率可以得知，浮誇新聞與非正常死亡為正向關係且大部分樣本點皆分佈在迴歸線附近，而少數饑荒災情相對嚴重的地區（如四川、安徽、青海）則被模型嚴重低估。

表12 分析結果

DV: 人均非正常死亡 (%)		$R^2: 0.239$			F-statistic: 8.143	
N: 28	Coef	Std coef	Std err	t	顯著性	
const	-0.056	-	0.201	-0.277	0.784	
浮誇新聞 (%)	0.04	0.488	0.188	2.854	0.008**	



資料來源：本研究自行繪製

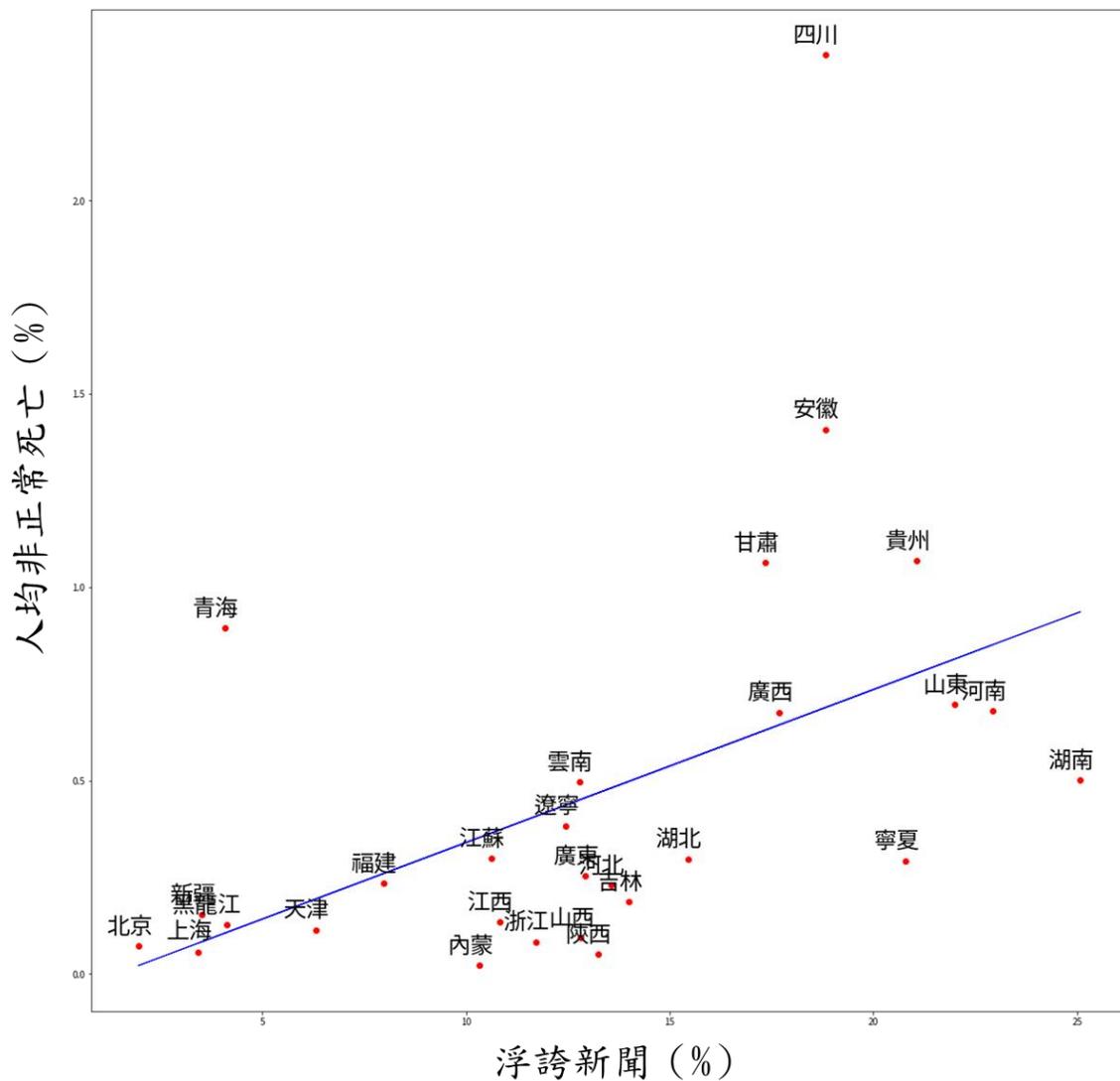


圖 15 迴歸線&散佈圖

資料來源：本研究自行繪製



有關於四川、安徽、甘肅、青海、貴州等饑荒較嚴重的省，本研究根據楊繼繩（2009）所撰之《墓碑》一書所記載的中國不同省份於大饑荒年代的真實情況來解釋這種現象。首先四川省是全國飢餓時間最長的省份，同時也是死最多人的省份，「三年大饑荒」爆發的時間點為 1959 年至 1961 年，但四川的饑荒是始於 1958 年冬季，結束於 1962 年秋季，亦即四川比其他地區多受饑荒摧殘了一年。當時四川被稱為天府之國，物產豐饒，然而中央誤以為當時四川的糧產仍然充裕，因此在饑荒年間國家糧食部門在四川召開全國糧食廳長會議，會中決定“四川每年都要調出大量糧食支援全國各地”。而當時四川的官員受於政治壓力，明知四川存糧已經供不應求卻仍答應中央的要求且不斷的釋放四川糧食充沛的消息。這也導致從 1957 年到 1960 年，四川是調出糧食多的省份，當時的中央仍認為四川糧食充足且能調出支援都會地區，這也導致當地輿論逐漸轉為“寧肯四川餓死人，也不能讓京、津、滬餓死人”亦即群眾認為中央想犧牲四川來保全北京、天津、上海等大都市的人民。1961 年秋後，多數省份已經開始走出饑荒，但一直到 1962 年四川還有大量饑荒災害的相關報告。當年四川的饑荒嚴重程度與其他省分不可比擬，過度的饑荒讓許多人民公社產生多起人吃人的案例，如大邑安仁區、灌縣蒲陽公社、灌縣崇義公社等。因此本研究認為，這種極端行為的歷史背景就是四川不論在糧食徵購數量或饑荒死亡人數遠超其他地區且迴歸模型嚴重低估非正常死亡與糧食徵購的主要原因。安徽省的饑荒非正常死亡人數僅次於四川，為全國第二多人死亡的省份。其歷史背景也與四川類似，安徽的饑荒是從 1958 年冬天開始，並在 1959 年春天達到高峰。與四川相似的地方是，當時安徽為毛澤東寄予厚望的重點城市，因此中央始終認為安徽省不可能有缺糧問題。然而安徽當時確實有大量飢餓死亡事件頻傳，對此安徽省委員認為，其原因並非因糧食不足，也不是國家徵購任務過重導致，而是人民的思想出了問題。中央亦認為安徽的饑荒問題是因為右傾保守思想的人瞞產私分所導致，且猜測有右傾思想的幹部帶頭私分糧食。並提出需要加強共產主義思想的宣傳，加大公社化的規模，最終導致



饑荒更為加劇。另外，楊繼繩（2009）在研究中也提到了梁志遠的著作：《關於“特種案件”的匯報——安徽亳縣人吃人見聞錄》，發現安徽省人吃人的案件比四川還多，書中記錄著當時吃人肉的行為非常盛行，市場上許多的肉類其實都是以人肉混充，且情況嚴重到政法單位成立專門破獲該類“特種案件”的部門。但由於中央的壓力，通常這類案件後續處理都是先嚴後寬，之後就不了了之。因此從上述案例可以看出當時安徽的災情也是十分嚴重。當然，本研究只深入探討了四川與安徽這兩個最嚴重地區的歷史背景，而其他饑荒相對嚴重的地區（甘肅、青海、貴州）其實也存在類似情況，只是沒有像四川與安徽的情況一樣極端，除了幹部虛報糧產導致一連串的徵購任務外，還有其他途徑會導致非正常死亡人數增加，例如地區的政治狂熱若特別嚴重，幹部壓榨農民的手段就越激烈，或是該省被中央視為重點關注地區，則中央會針對當地加強人民公社化的施政規模。因此本研究認為這些特殊的影響途徑是導致迴歸模型低估主因。

第五章 結論

第一節 主要研究發現



壹、浮誇新聞的表述方式

由關鍵字分析的結果可以分析出，《人民日報》所報導的糧食浮誇新聞主要是利用大豐收、糧食充足來結合「人民公社化」運動，建立出人民公社化空前的成功，舉國上下因為政策施行順利帶來了一遍繁華的假象。這類新聞的內容是地方幹部因為壓力而上報假產量後再經過《人民日報》的添油加醋後演化而成。其中又可以將這類浮誇新聞再細分為兩類，一類為簡要形容全國的高產盛況，其主要目的在於建立「大躍進」政策順利進行的假象，是一種《人民日報》對中共中央展現忠誠度的表現。另一類則是描述某地區的高產盛況、生產成果等成功案例。這類新聞的目的主要是透過大力報導某個地區的成果來增加其他地區幹部的壓力，進而促進產量提升。例如《人民日報》曾在大躍進期間發佈《比一比》專欄，公布各省的生產紀錄與目標完成進度（姚桂榮，2012）。

此外，形容各地豐收的字詞方面，《人民日報》除了使用「豐收」、「高產」、「破紀錄」等形容詞外，還會使用「萬畝」、「萬斤」、「千萬噸」等巨大單位的量詞來展現豐產的盛況。整體來說糧食浮誇新聞內的元素包羅萬象，但其內容主要離不開「糧食高產發生的地區」、「誇大的形容詞」、「巨大單位的生產成果」、「高產的作物種類」以及「對中央政策的歌功頌德」。另外，本研究亦發現浮誇關鍵字的聲量從 1958 年開始攀升，一直到 1960 年到達最高峰，因此可以得知當年全國的浮誇風氣從大躍進開始後不斷氾濫，最後到大饑荒後期才開始快速收斂。透過關鍵字分析與聲量分析我們不僅瞭解了當時浮誇新聞組成的脈絡與概念架構，也可以得知浮誇訊息會隨著「大躍進」的演進而不斷增加。這些新聞的目的不僅是為了激勵各地幹部的生產進度，還會順帶傳達黨的思想，以及強化民眾對中央

政策的認同感。



貳、浮誇新聞與饑荒嚴重程度之關係

從浮誇新聞百分比與人均非正常死亡百分比之盒鬚圖可以看出，不同省份的浮誇新聞樣本中不存在離群值，但非正常死亡卻存在四川的離群值，代表著四川之饑荒嚴重程度與其他地區有非常大的落差，且深入觀察數據可以發現各省饑荒發生的年度也不盡相同，較發達的地區如北京、天津、上海等地區或是新疆、內蒙等偏遠地區的饑荒情況則不嚴重。另外，以全國的角度來看，浮誇新聞與非正常死亡在「大躍進」五年期間的趨勢漲跌普遍一致，皆在 1959 年開始攀升，直至 1960 年達到高峰後開始下降，此現象與「大躍進」政策之發展進程一致。

本研究透過迴歸分析驗證了：「浮誇新聞會正向影響饑荒死亡人數」，但此結果可能不適用於饑荒特別嚴重的地區。迴歸線散佈圖中可以發現，如四川、安徽、青海與甘肅等距離迴歸預測線較為遙遠，在深入瞭解這些地區的歷史背景與當年的政策後可以發現，造成這些地區饑荒如此嚴重的原因除了浮誇新聞以外還受其他更重要的因子或途徑（如中央的關注程度、幹部的政治狂熱程度、生產任務的難易度等、是否為農業重鎮）所影響。

第二節 後續研究建議

本研究經過對《人民日報》文本進行一系列的探究與檢定，成功驗證了研究假設。然而這項論點僅為影響饑荒嚴重程度的眾多因素之一，若要更全面了解饑荒的成因與後續影響的其他因素，除了需尋找與「大躍進」相關且「可信度高」的資料進行量化分析外，還必須輔以質化研究來佐證其因果關係。由於中國官方常會因政治問題而透過封鎖或竄改對其政權不利之歷史數據，藉此來隱瞞真相。故本研究也難以在現有資料上找到其他可靠的數據進行實驗，這同時也是本研究面臨的最大限制。



本研究使用《人民日報》所發佈的浮誇新聞來衡量饑荒嚴重程度的關聯，雖然可證明出兩者之間具有關聯性，但實際上浮誇新聞亦可能是透過其他中介變數來影響饑荒，而本研究僅用「浮誇新聞導致中央誤信而增加生產任務，進而導致饑荒更嚴重」的邏輯來解釋結論，其實有解釋能力不足的疑慮。另外，本研究礙於資料取得難易度問題而以省級行政區作為分析單位進行分析，但其實以省為單位進行分析可能忽略縣級行政區、鄉級行政區之間的差異。因此本研究建議後續對此議題有興趣的研究者可以嘗試尋找縣級、鄉級行政區的地理資料，或尋找適合的中介變量來嘗試對饑荒嚴重程度進行因果推論。在研究方法之選擇方面，過往進行相關研究的學者主要會因為資料取得難易度高而選擇以文獻分析或深度訪談等質化方法進行研究。然而質化研究方法主觀性相對較高，因此在此類需客觀事實來佐證的議題仍建議以量化研究方法來進行相對合適。以迴歸分析來說，由於模型本身具有因果關係的設定，因此能夠較為客觀的呈現事實證據，在相關數據難以取得之時通常會利用文獻理論嘗試定義一個新的衡量指標來解決。解釋變數方面，使用「大躍進」時期各地區的平均家戶所得、各地區的糧食配給數，或者建立新的指標，例如不同地區的都市化程度、政治狂熱程度、真實糧食生產能力、農業現代化程度等，皆可能對饑荒的嚴重程度有解釋能力。另外，本研究因時間與資料取得難易度的限制，只能選擇以饑荒非正常死亡人數作為衡量饑荒嚴重程度的變量。然而饑荒的嚴重程度其實可以從其他面向切入衡量，例如該期間的生育減少人數、浮腫病患病人數、失蹤人口數等等。若以文字探勘方法來進行量化分析，本研究則建議可以放大檢視一些饑荒情況較嚴重的地區，並以地區媒體發佈的新聞作為切入點進行研究，由於地區的新聞並非站在全國與中央的角度進行報導，而是只反映該地區當前的情況，因此有助於得知該地區在大躍進時期更詳細的資訊。透過分析少數地區的情況、尋找新的變量，或是從其他文本切入分析，皆是後續梳理大饑荒全貌可以進一步探討的方向。

參考文獻



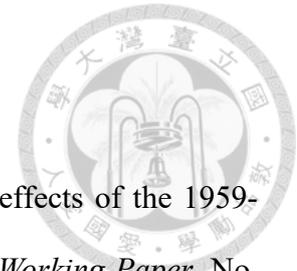
壹、中文

- 中共中央文獻研究室 (2009)。《毛澤東文集，第五卷》。北京：人民出版社。
- 中國國家統計局 (1984)。《中國統計年鑑》。北京：中國統計出版社。
- 中國國家統計局綜合司 (1964)。《糧食部計劃司糧食資料提要》。北京：中國統計出版社。
- 方雪琴 (2004)。〈資訊公開與媒體理性—試論危機傳播中的輿論引導策略〉，《中州學刊》6：183-185。
- 王官德、劉承宗、李化成 (2008)。《中國共產黨史》。臺北：五南圖書出版。
- 王毓莉 (1998)。〈中共改革開放政策對電視事業經營之影響〉，《新聞學研究》57：27-49。
- 王毓莉 (2009)。〈中國大陸「南方週末」跨地區新聞輿論監督報導之研究〉，《新聞學研究》100：137-186。
- 吳柳林 (2011)。〈大躍進時期輿論場的形成及原因分析〉，《今傳媒》1：143-144。
- 杜聖聰 (2008)。《兩岸真相密碼：中共對台宣傳的政策、作為與途徑》。台北：秀威。
- 邵軒磊、曾元顯 (2018)。〈文字探勘技術輔助主題分析—以「中國大陸研究」期刊為例〉，《問題與研究》57(1)：29-62。
- 姚桂榮 (2012)。〈“大躍進”期間的新聞傳播及其對從眾行為的影響〉，《史學月刊》12：129-133。
- 韋旭 (2008)。〈主流媒體輿論與公眾輿論的和諧關係建構〉，《廣西大學學報》30(2)：138-141。
- 茹倩 (2013)。〈“大躍進”中新聞傳播及原因分析〉，《新聞與傳播》12：276-277。

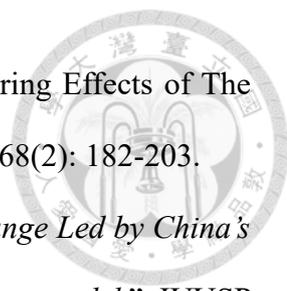


- 曹樹基 (2005)。《大饑荒：1959-1961 年的中國人口》。香港：時代國際。
- 陳永發 (2003)。《中國共產革命七十年》。臺北：聯經。
- 陳福成 (2000)。《國家安全與戰略關係》。台北：時英出版社。
- 黃永鋒、李代友、韓華淑 (2015)。〈論《人民日報》對“大躍進”運動的影響〉，
《傳承》7：22-24。
- 楊繼繩 (2009)。《墓碑——中國六十年代大饑荒紀實》。香港：天地圖書。
- 葛玲 (2010)。〈“寧左勿右”的政治行動邏輯——皖西北臨泉縣浮誇風中的幹部行為研究〉，《當代世界社會主義問題》4：3-17。
- 董芹 (2007)。〈大躍進中《人民日報》傳媒功能的缺失〉，《青年記者》20(10)：
90。
- 董傳嶺 (2015)。〈1959 ~ 1961 年饑荒成因探析——以山東省梁山縣狀況為例〉，
《安徽農業科學》2：389-392。
- 靖鳴、劉銳 (2008)。〈1958 年“大躍進”期間浮誇新聞的成因及其啟示〉，《視聽》
11：10-15。
- 樑柱 (2013)。〈毛澤東《論十大關係》的方法論意義〉，《高校理論戰線》12：42-
47。
- 潘忠黨、陳力丹、李良榮、趙月枝、孫旭培、吳飛、魏永征、單波、黃旦、徐
貴、李霞、陳衛星 (2008)。〈反思與展望：中國傳媒改革開放三十周年筆
談〉，《傳播與社會學刊》6：17-48。
- 蔣正華、李南 (1998)。〈中國人口動態估計的方法與結果〉，《中國人口年鑑》：
94-106。
- 閻茂旭、劉璿 (2006)。〈對“雙百”方針受挫的歷史反思〉，《淮陽工學院學報》
15(6)：50-52。
- 鐘萬全 (2003)。〈『浮誇風』導致大饑荒紀實〉，《廣東黨史》5：21-26

貳、英文



- Almond, D., Edlund, L., Li, H. and Zhang, J. (2007). "Long-term effects of the 1959-1961 China famine: mainland China and Hong Kong." *NBER Working Paper*, No. 13384.
- Banister, J. (1991). *China's Changing Population 1st*. Stanford University Press.
- Bengtsson, E (2017). *There is always a little famine going on - A study of articles in the People's Daily during the Great famine in China*. Master Degree Project, Lund University Language and Linguistics.
- Calot, G. (1985). "The suggestion to analysis the Chinese census data." IUUSP international population conference, Florence, Italy, June 5.
- Coale, A. (1984). *Rapid Population Change in China, 1952-1982*. Washington, DC: The National Academies Press.
- Dikötter, F. (2010). *Mao's Great Famine*. London: Bloomsbury Press.
- Garnaut, A. (2013). "Hard facts and half-truths: The new archival history of China's Great Famine." *China Information*, 27(2):223-246.
- Guan, T (2018). "Framing the Boundary of Sino-Japanese Conflicts in China's Communication Sphere: A Content Analysis of the News Coverage of Japan and Sino-Japanese Controversies by the People's Daily between 2001 and 2015." *Journal of Chinese Political Science*, 23(3): 603-618.
- He, Q. (2012). "Financial deregulation, credit allocation across sectors, and economic growth: evidence from China." *Journal of Economic Policy Reform*, 15(4): 281-99.
- He, Q. (2013). "Price scissors and economic growth: the role of openness." *Pacific Economic Review*, 18: 60-78.

- 
- He, Q. and Sun, M. (2016). “Central Planning Legacies: The Lingering Effects of The Great Leap Forward in China.” *Bulletin of Economic Research*, 68(2): 182-203.
- Hong, M., Yang, M and Ning, J. (2013). “A Study of Population Change Led by China’s Great Famine (1958-1961) —based on modified Lee-Carter model.” IUUSP international population conference, Busan, August 26.
- Houser, D., Sands, B. and Xiao, E. (2009). “Three parts natural, seven parts man-made: Bayesian analysis of China’s Great Leap Forward demographic disaster.” *Journal of Economic Behavior & Organization*, 69(2): 148-159.
- Hsiung, P. C. and Wang, Y. (2018). “Unmasking China’s Great Leap Forward and Great Famine (1958-1962) Through Shunkouliu.” *Qualitative Inquiry*, 25(8): 1-11.
- Kang, A., Ren, L., Hua, C., Dong, M., Fang, Z and Zhu, M. (2021). “Stakeholders’ views towards plastic restriction policy in China: Based on text mining of media text.” *Waste Management*, 136: 36-46.
- Karlsson, K. and Schoenhals, M. (2008). *Crimes against humanity under communist regimes - Research review*. Lund University: Forum for Living History.
- Kirby, E. S. (2011). “China's first five-year plan.” *Journal of The Royal Central Asian Society*, 42(4): 269-274.
- Le, Q. and Mikolov, T. (2014). “Distributed Representations of Sentences and Documents.” PMLR International Conference on Machine Learning, Beijing, China, June 21.
- Meng, X., Qian, N. and Yared, P. (2015). “The Institutional Causes of China’s Greatest Famine, 1959-1961.” *Review of Economic Studies*, 82: 1568-1611.
- Mikolov, T., Chen, K., Corrado, G and Dean, J. (2013). “Efficient Estimation of Word Representations in Vector Space.” ICLR 2013 conference, Arizona, January 17.
- Nassirtoussi, A.K., Aghabozorgi, S., Wah, T.Y. and Ngo, D. C. L. (2014). “Text mining



- for market prediction: A systematic review.” *Expert Systems with Applications*, 41(16): 7653-7670.
- Ning, Z., Yuefeng, L and Sheng, T. W. (2012). “Effective Pattern Discovery for Text Mining.” *IEEE Transactions on Knowledge and Data Engineering*. October 28.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Aidan, N., Gomez, L. K. and Polosukhin, I. (2017). “Attention Is All You Need.” *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, online, December 22.
- Vreese, C. H. (2005). “News framing: Theory and typology.” *Information Design Journal*, 13(1): 51-62.
- Xu, H., Li, L., Zhang, Z. and Liu, J. (2016). “Is Natural Experiment a Cure? Re-examining the Long-Term Health Effects of China’s 1959–1961 Famine.” *Social Science and Medicine*, 148:110-22.
- Xu, H. and Tian, G. (2020). “Is Lying Contagious? Spatial Diffusion of High-Yield “Satellites” during China’s Great Leap Forward.” *American Journal of Sociology*, 126(3): 632-672.
- Xu, J., Hu, Y. and Miyoshi, E. (2018). “A Consideration of Media Environment Regarding Air Pollution Problems in China: Based on the Content Analysis of the Reports of “People’s Daily” from Jan. 1,1970 to Nov. 30,2011.” 2018 International Joint Conference on Information, Media and Engineering (ICIME), Osaka, Japan, January 14.
- Yan, X. Y. (2020). “The Impact of social media on Traditional Mainstream Media— A Case Study of People’s Daily.” International Seminar on Education, Management and Social Sciences (ISEMSS 2020), Beijing, July 17.
- Zhang, S. Q. (2020). “Analysis on the Noise of People’s Daily Information Expression

During Covid-19 Epidemic.” Proceedings of the 2nd International Conference on Literature, Art and Human Development (ICLAHD 2020). Xiamen, December 17.

